University College London

The Role of Serum Amyloid P in Amyloidosis

By Rebecca Coker

A Thesis Submitted for the Degree of Doctor of Philosophy

2018

UCL Doctorate in Biomedical Sciences

Thesis declaration form

I confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signature: RCoker

Name: Rebecca Coker

Date:4/9/2018

Abstract

Amyloid is a term used to describe a set of diseases caused by protein misfolding. There are 36 proteins known to undergo amyloidosis to date. They each unfold from their unique native states and refold into a specific cross β-sheet motif to which various macromolecules can bind. These protein deposits invade tissues and organs disrupting tissue structure and function.

The protein serum amyloid P (SAP) is a universal component to all amyloid deposits regardless of the protein involved. It is a disc shaped protein which is known to protect amyloid fibres against degradation. The role of SAP in amyloidosis has been a subject of discussion for many years. Some have found SAP to enhance fibre formation whereas others have found SAP to prevent it.

This thesis can be divided into two parts. The first half focuses on the molecular analysis of SAP in relation to amyloid fibre growth whereas the second half focuses solely on the structural aspects of SAP.

Chapters 2 and 3 demonstrate that SAP actually exhibits a dual role in relation to fibrillogenesis. It is capable of acting as both a chaperone and a fibrillogenic enhancer depending on its conformation at the time. As a monomer SAP is responsible for the enhancement and stabilisation of fibres whereas in its dimerised form it demonstrates characteristics similar to that of a molecular chaperone.

Recent developments in amyloid treatment involve the removal of SAP from deposits using anti-SAP antibodies. As of yet, structural analysis has not been carried out on the SAP:antibody complex. Chapters 4 and 5 conclude this thesis by using crystallographic techniques to demonstrate that a single antibody binds to an SAP molecule during dissociation from amyloid deposits.

Impact Statement

Although amyloid diseases are considered rare, they are likely more common than previously thought. The wide range of symptoms associated with amyloid disease often leads to misdiagnosis and thus there is a lack of literature regarding this category of diseases. The field has been grossly under researched as a result of this. So far 36 proteins have been associated with amyloid diseases and this number is continually increasing. Each of these diseases requires unique treatment to combat its unique set of symptoms and any delay can easily result in irreversible if not fatal consequences. Not enough is known about amyloid in general let alone each specific type to allow for effective, individual treatment plans. Because of this, the idea of having a general SAP treatment plan is extremely appealing.

This thesis shows that the serum protein, SAP, has the ability to reduce *or* enhance amyloid fibre mass depending on its conformation.

By showing its ability to reduce amyloid fibre mass, the possibility of using SAP decamers as a general treatment for amyloid can be taken into consideration. This removes the need for individual treatment of each of the 36 established diseases. Not only is SAP universal to all forms of amyloid, it is also a component already present in the human body so is unlikely to impose further damage to the host.

Contents

Tables

Figures

Abbreviations

AA	-	Amyloid A Amyloidosis
Aβ	-	Amyloid Beta
Aβ2M	-	β-microglobulin Amyloidosis
ADDLs	-	Aβ-derived diffusible ligands
ADP	-	Adenosine Diphosphate
AL	-	Light Chain Amyloidosis
APH-1	-	Anterior Parynx-Defective 1
APP	-	Amyloid Precursor Protein
APTG	-	*p*-amino-phenyl-β-D-thio-galactosidase
ATP	-	Adenosine Triphosphate
ATTR	-	Transthyretin Amyloidosis
BACE	-	Beta-Site APP –Cleaving Enzyme
BCA	-	Bicinchoninic Acid
CA	-	Carbonic Anhydrase
CBP	-	chitin binding protein
CCD	-	Charged Coupled Device
CD	-	Circular Dichroism
CDR	-	Compliment Determining Region
C_H	-	Constant Heavy Domain
C_L	-	Constant Light Domain
CPHPC	-	R-1-[6-[R-2-carboxy-pyrrolidin-1-yl]-6-oxo-hexanoyl] pyrrolidine-2-carboxylic acid
CPL	-	Circularly Polarised Light
CSF	-	Cerebrospinal Fluid
DMSO	-	Dimethyl Sulfoxide
EDTA	-	Ethylenediaminetetraacetic Acid
EGTA	-	Ethylene Glycol Tetra acetic Acid
EM	-	Electron Microscopy
E.coli	-	*Escherichia coli*
Fab	-	Fragment Antigen Binding

FAC	-	Familial Amyloid Cardiomyopathy
FAP	-	Familial Amyloid Polyneuropathy
Fc	-	Fragment Crystalline
FTIR	-	Fourier Transform Infrared Spectroscopy
G6P	-	Glucose-6-phosphate
GFP	-	Green Fluorescent Protein
GSK	-	GlaxoSmithKline
GST	-	glutathione-S-transferase
His-tag	-	Polyhistidine-Tag
HPLC	-	High Performance Liquid Chromatography
HRP	-	Horseradish Peroxidise
Ig	-	Immunoglobulin
IgA	-	Immunoglobulin A / Immunoglobulin α
IgD	-	Immunoglobulin D / Immunoglobulin δ
IgE	-	Immunoglobulin E / Immunoglobulin ε
IgG	-	Immunoglobulin G / Immunoglobulin γ
IgM	-	Immunoglobulin M / Immunoglobulin μ
LB	-	Luria Broth
LLG	-	Log Likelihood Gain
LRP	-	Low density lipoprotein receptor-related protein 1
MAD	-	Multi-wavelength Anomalous Diffraction
MBP	-	Maltose Binding Protein
MPD	-	2-Methyl-2,4-pentanedio
mRNA	-	Messenger RNA
NHS	-	N-hydroxysuccinimide
NMDA	-	*N*-methyl-D-aspartate
NMR	-	Nuclear Magnetic Resonance
PBS	-	Phosphate-Buffered Saline
PDB	-	Protein Data Bank
PC	-	Proprotein convertase
PCB	-	Polychlorinated Biphenyl
PEG	-	Polyethylene Glycol
PNS	-	Peripheral Nervous System
PrP	-	Prion Protein

PSEN	-	Presenilin
PVDF	-	Polyvinylidene Difluoride
RER	-	Rough Endoplasmic Reticulum
SAD	-	Single-wavelength Anomalous Diffraction
SAP	-	Serum Amyloid P
SDS	-	Sodium Dodecyl Sulfate
SDS PAGE	-	Sodium Dodecyl Sulfate Polyacrylamide
SLE	-	Systemic Lupus Erythematosus
SSA	-	Senile Systemic Amyloidosis
TBS	-	Tris Buffered Saline
TEM	-	Transmission Electron Microscopy
ThT	-	Thioflavin T
TRIS	-	Tris(hydroxymethyl)aminomethane
tRNA	-	Transfer RNA
TTR	-	Transthyretin
V_H	-	Variable Heavy Domain
V_L	-	Variable Light Domain

Buffer Solutions

Alkaline buffer	-	100mM TRIS, 5mM EDTA, pH10.
Coomassie	-	0.5g Brilliant blue R, 50ml Aceitic acid, 225ml dH_2O, 225ml Methanol.
De-stain	-	10% acetic acid, 10% methanol, 80% distilled water.
Digestion buffer	-	20mM sodium phosphate, 20mM Cysteine, 0.01% Na Azide, pH 7.
Laemmli	-	10ml, 0.25M TRIS pH 6.8, 4ml glycerol, 4ml water, 0.8g SDS, 2ml β-mercapto ethanol, 8mg Bromo-phenyl blue.

Overnight Expression Buffer:

- 1ml 1M $MgSO_4$,
- 20ml “50x5052” – 25% w/v glycerol, 13.9mM glucose, 29.2mM α-lactose.
- 50ml “NPS” – 0.5M $(NH_4)_2$ SO_4, 1M KH_2PO_4, 1M Na_2HPO_4.
- LB broth to 1 litre,
- 30µg/ml chloramphenicol,
- 30µg/ml ampicillin.

PBS Buffer	-	0.14M NaCl, 3mM KCl, 10mM Na_2HPO_4 2mM KH_2PO_4 pH 7.4.
Purification Buffer (made on day of use)	-	10mM Tris-HCL, 1mM EDTA, pH8.5.

Running Buffer (5x)	-	125mM TRIS, 1M Glycine, 0.5% w/v SDS, pH 8.3.
Sample Buffer	-	20mM sodium phosphate, 5mM EDTA, 0.01% Na Azide, pH 7.
Separating gel (12%)	-	3ml 30% bis-Acrylamide (protogel), 1.87ml 4 x TRIS, 0.4% w/v SDS, pH 8.8 buffer, 2.6ml distilled water, 50µl 10% w/v APS, 10µl TEMED.
Sodium Acetate Buffer (Tajiri)	-	0.2M NaOAc, 0.2M HOAc, pH 5,6 and 7
Sodium Acetate Buffer (Kugimiya)	-	20 mM NaOAc, 100 mM NaCl, pH 3.0
Stacking gel	-	0.65ml 30% bis-Acrylamide (protogel), 1.25ml 4 x TRIS, 0.4% w/v SDS, pH 6.8 buffer, 3.05ml distilled water, 50µl 10% w/v APS, 10µl TEMED.
TC Buffer	-	20mM TRIS, 600mM NaCl, 2mM CaCl, 0.01% Na Azide, pH8.
TE Buffer	-	20mM TRIS, 150 NaCl, 0.01% Na Azide, pH8.
THS Buffer	-	20mM TRIS, 600mM NaCl, 0.01% Na Azide, pH8.
(n)M Urea Buffer pH8.5.	-	(n)M Urea, 10mM Tris-HCL, 1mM EDTA,

Chapter 1

An Introduction to Proteins, Amyloid and SAP

An Introduction to Proteins

Proteins are the building blocks of life. They perform a vast array of duties within organisms and are responsible for the structure, function and regulation of the body's tissue and organs. In fact, proteins are responsible for nearly every task of cellular life. They receive signals from outside the cell, organise intracellular responses and are accountable for cell shape, product manufacture and waste clean-up.

They are the work-force molecules of the cell and are as diverse as the functions they serve. For instance, some proteins act as catalytic enzymes for biochemical reactions in the body whereas others have more of a structural or mechanical role. Take for an example DNA polymerases and keratin. DNA polymerases are enzymes responsible for synthesising DNA molecules during replication. In contrast, keratin is a protein in your hair, skin and nails which is responsible for maintaining cell shape and forming a protective layer around the body.

Proteins can be big, small, single or multi-unit complexes. They can exist as mostly hydrophobic or mostly hydrophilic structures and also have the ability to change their shape if needed. All of these differences arise from the unique amino acid sequences that make up each protein (Lesk, 2010).

Amino Acids

Every protein is a macromolecular chain made up of hundreds or thousands of amino acids - a simple organic compound containing both a carboxyl (—COOH) and an amino (—NH_2) group. The chain is synthesised by the ribosome complex in a process referred to as DNA translation. Here, each amino acid is covalently bound to its neighbour via a condensation reaction between the carboxyl group of one and the amine group of another. Water is released and a CO-NH bond – *a.k.a.* a peptide bond – is produced. Once

linked in the protein chain, an individual amino acid is referred to as a residue and the linked series is often referred to as the backbone primary sequence (Berg, Tymoczko, and Stryer 2002).

There are 20 different amino acids present in nature, all of which share three common structural features: The first being that they each have a central α-carbon to which a hydrogen atom, amino group, carboxyl group and a variable side chain are bound. It is these 20 variable side chains which give the amino acids unique properties and characteristics (more detail to follow). The amino acid 'proline' is the only example which differs from this basic format as its variable side chain links back to the nitrogen of the main chain removing its ability to acts as a hydrogen bond donor (Lodish, Berk, and Kaiser 2007). The second structural feature is that all amino acids are chiral *i.e.* their mirror image cannot be superimposed; glycine, due to its variable side chain consisting of another lone hydrogen atom, is the only exception to this rule. Finally, the α-carbon is asymmetric, meaning that there are two different ways in which groups can be arranged around it. This allows for two isomers to form, each of which interacts differently with polarised light. The predominant isomer is the L-form although the D-form is present in bacterial cell walls and some antibiotics such as valinomycin and gramicidin A. Only the L-amino acids exist in natural proteins with a few exceptions following post-translational modifications(Lesk, 2010) **(Figure 1) (Figure 2)**.

Protein Structure

Proteins are dynamic molecules whose functions almost invariably rely on interactions with other molecules. Those molecules specific to proteins are referred to as ligands. Ligands have a variety of roles and structures in terms of protein binding. They can be small molecules, ions or even other protein complexes. Some act as substrates, were as others behave as inhibitors, activators, and even neurotransmitters. Protein-ligand binding is typically a reversible interaction, often results in a change of conformation of

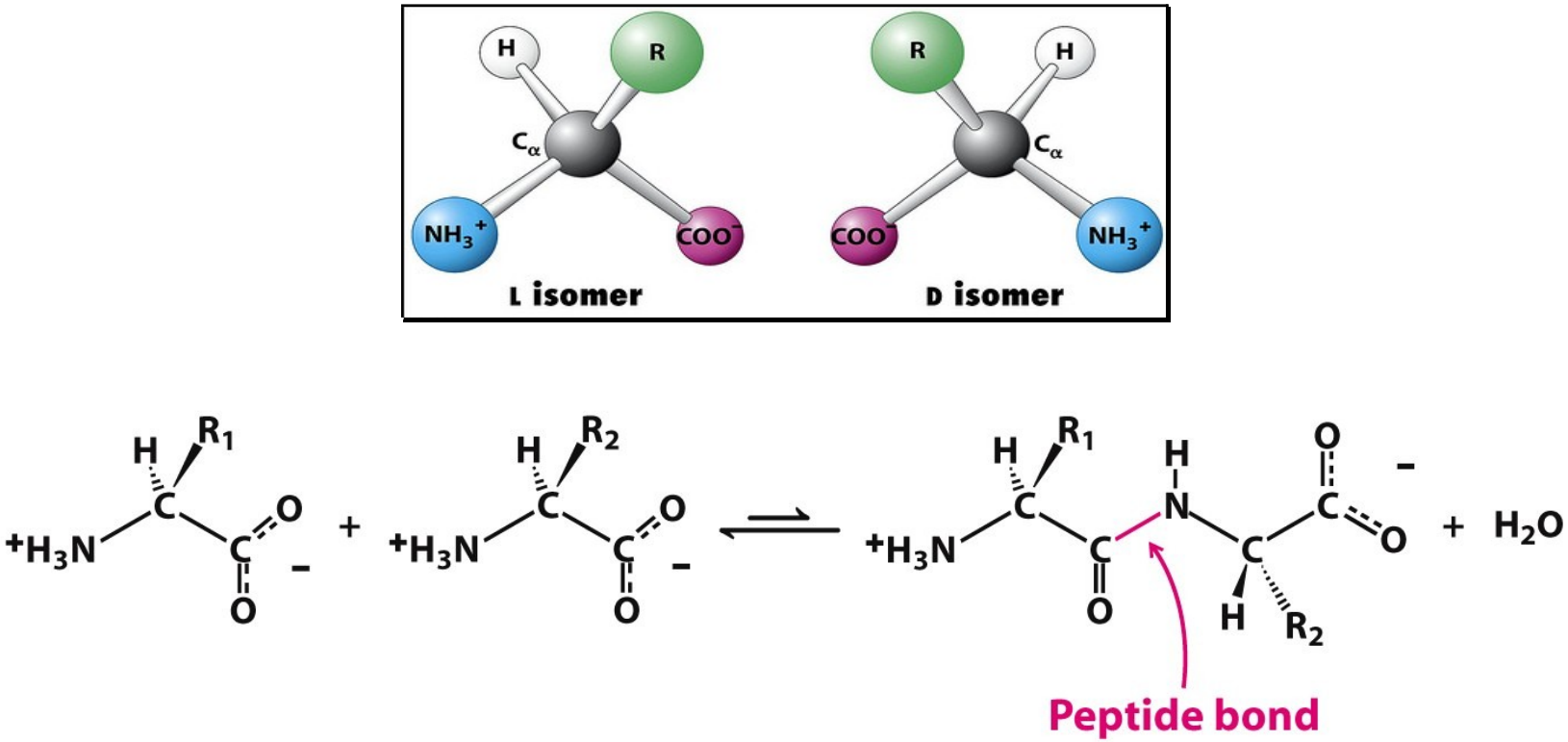


Figure 1: (Above) Comparison of the two amino acid isomers. Each isomer interacts differently with polarised light. The predominant form is the L form although the D form is present in bacterial cell walls and some antibiotics such as valinomycin and gramicidin. Only L-form amino acids exist in natural proteins with a few exceptions post-translational modifications. (Below) The formation of a peptide bond between two amino acids. A peptide bond takes place between two adjacent amino acids in a polypeptide chain. It is a condensation reaction which occurs when the carboxyl group of one molecule reacts with the amino group of the other molecule, releasing a molecule of water (H2O) (Berg, Tymoczko, and Stryer 2012).

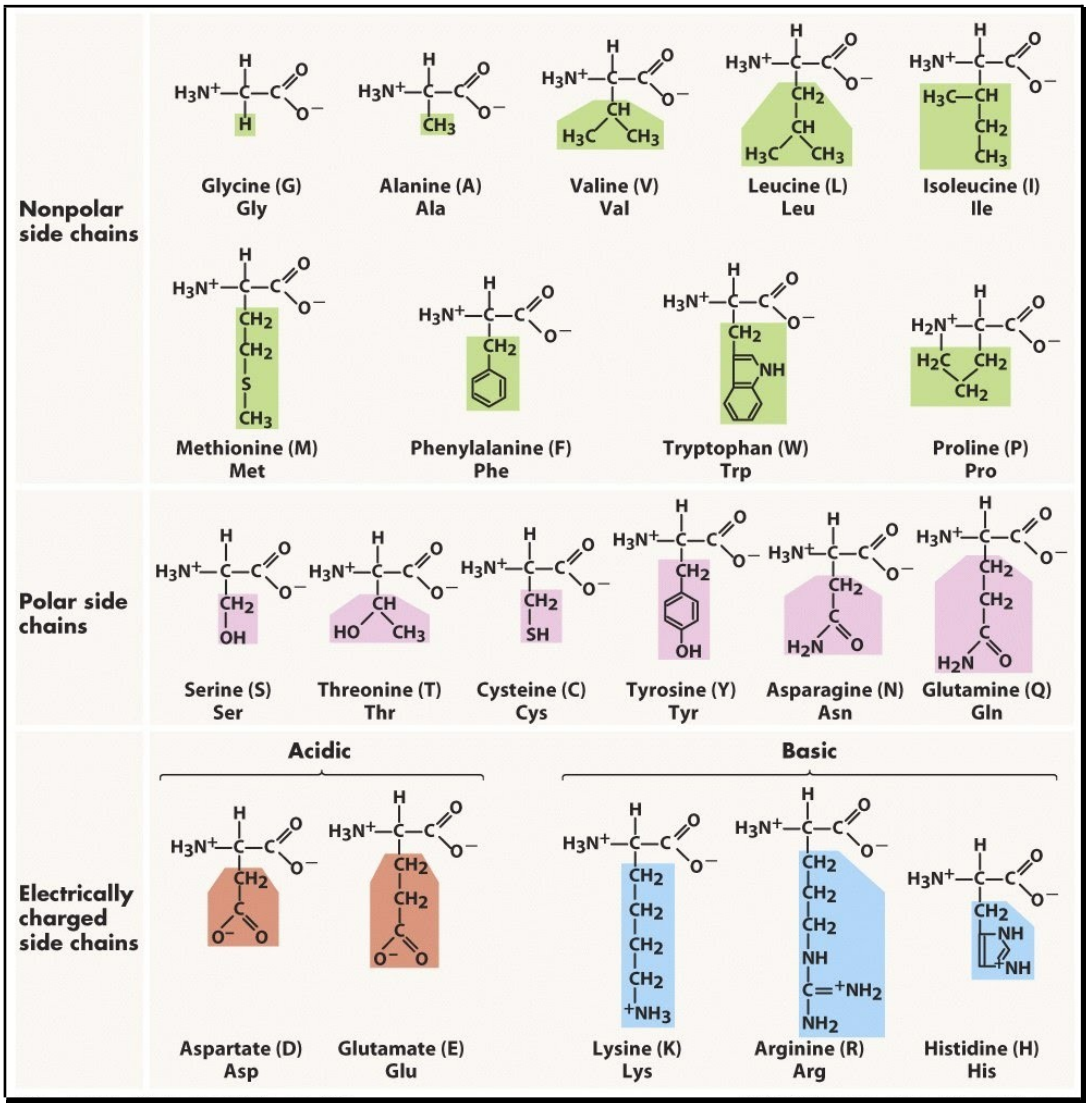


Figure 2: The structure, full names, single letter code and 3 letter codes of the 20 amino acids at a physiological pH of 7.4. The variable side chains of each are highlighted. The amino acids have been categorised into non polar, polar neutral, polar acidic and polar basic subgroups. The non polar amino acids are composed purely of hydrocarbon alkyl groups or benzene rings. The more alkyl groups present, the more non-polar the amino acid is i.e. valine is more non-polar than alanine and leucine is more non-polar than valine. Those which are polar but neutral behave as such due to the acid and amine groups on the α-carbon neutralising each other creating a "zwitterion". However, the polar acidic and basic side chains each have an additional acid/amine group producing a net effect of either positive or negative charge respectively (Lodish, Berk, and Kaiser 2007).

the target protein and usually involves some form of ligand mediated signal transmission.

Ligand-mediated signal transmission is essential to all life processes, however, in order for interactions to actually occur, the binding site of the protein must be complimentary to the specific shape and charge of its corresponding ligand; *i.e*. binding is successful when the protrusions, clefts and grooves of the protein binding site match those of the ligand in a fashion similar to a lock and key. Binding occurs through a variety of intermolecular forces including ionic binding, hydrogen bonds and Van der Waals forces. The rate of binding is called affinity, and this measurement typifies a tendency or strength of the effect. Binding affinity is actualized not only by host-guest interactions, but also by solvent effects that can play a dominant, steric role which drives non-covalent binding in solution. The solvent provides a chemical environment for the ligand and receptor to adapt, and thus accept or reject each other as partners.

Ligands by nature are three dimensional structures. Therefore protein-ligand binding can only occur if proteins themselves also have a three dimensional shape. Every protein has a unique polypeptide sequence and thus folds into a unique globular shape. This does not occur spontaneously. Instead, a series of steps are followed in order to achieve optimum folding (Bolsover et al. 2003).

Protein folding can be categorised into four groups: primary, secondary, tertiary and quaternary structures. The primary structure refers to the sequential order of the amino acid residues in the polypeptide chain. The secondary structure refers to the folding of the chain into regularly repeating local formations and the tertiary form is the folding of those formations into a three dimensional globular shape. The final quaternary shape occurs when several protein molecules form a combined structure to function as a single protein complex **(Figure 3)**.

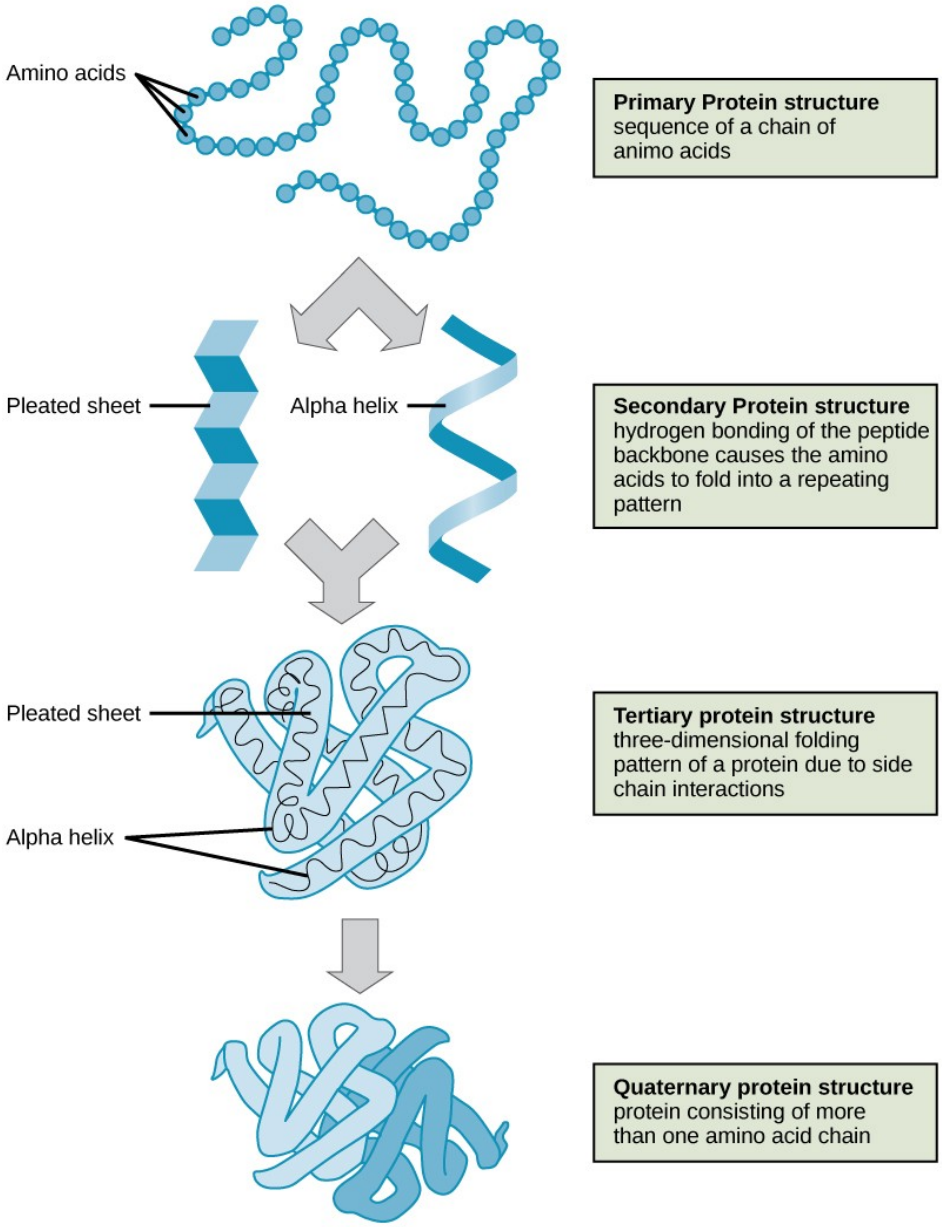


Figure 3: A diagram of the four protein folding categories: primary, secondary, tertiary and quaternary. The primary structure refers to the sequential order of the amino acid residues in the polypeptide chain. The secondary structure is the folding of the backbone into regularly repeating local formations (α-helices and β-pleated sheets) and the tertiary structure is the folding of those formations into a three dimensional globular shape. The quaternary structure occurs when several protein molecules form a combined structure to function as a single protein complex.

Primary Structure

The primary structure acts as the backbone of a protein. It spans from an un-bonded primary amine group (N-terminus) to an un-bonded carboxyl group (C -terminus) and has the variable side chains of each residue protruding in a perpendicular manner along its length. The actual folding of this backbone is influenced by steric constraints posed by the covalent peptide bond. The peptide bond is planar and is often described as having partial double bond characteristics. This is due to the fact that it has a bond length of 1.32Å instead of the expected 1.49Å or 1.27Å of a single (C-N) or double (C=N) carbon-nitrogen bond respectively. As a result of this, the peptide bond experiences restricted rotation around the axis of the carbon and nitrogen atoms which in turn restricts folding options for the backbone (Sipe, 2006).

Secondary Structure

Folding of the polypeptide backbone is not a case of random trial and error. In fact, secondary structure folding is influenced by two things. The first of these is the steric constraints presented by the peptide bond and the second is the charge and size of the variable side chains which surround it. With regards to the latter, it is important to know that there are a variety of chemical structures and properties which make up the side chains. These can be divided into two categories: polar and non polar. Within these two groups each amino acid can also be described as being acidic, basic or neutral.

The amino acids whose variable side chains include various functional groups such as acids, amides, alcohols, and amines are classified as polar. The level of polarity depends on the functional group itself and on the number of carbon-hydrogens present in any alkane or aromatic portion of the side chain. The greater the difference of electronegativity in a bond the more polar the bond is. For example, aspartic acid is more polar than serine as an acid

functional group is more polar than an alcohol one. However, serine is more polar than threonine since threonine has one more methyl group than serine. The methyl group gives more non-polar character to threonine.

Non polar amino acids are composed purely of hydrocarbon alkyl groups or benzene rings. The more alkyl groups present, the more non-polar the amino acid is. Examples include valine, alanine and leucine. As with the polar group, the number of alkyl groups influences the polarity. The more alkyl groups present, the more non-polar the amino acid will be. This in effect makes valine more non-polar than alanine and leucine more non-polar than valine.

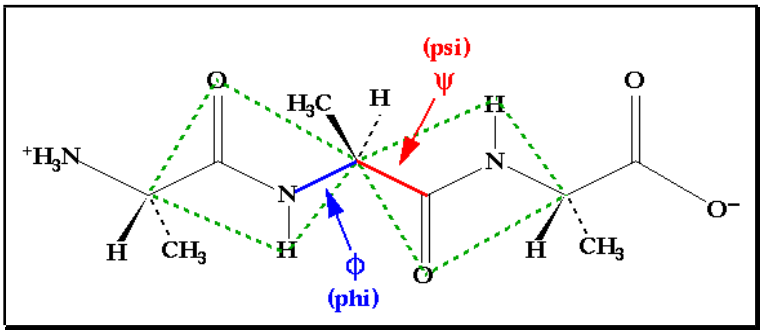
There are two amino acids which have acidic functional groups: aspartic acid and glutamic acid. Normally, an amino acid produces a nearly neutral solution since the acid group and the basic amine group on the α-carbon neutralize each other creating a “zwitterion”. However, the above examples each have an *additional* acid group producing a net acid effect. In order to create a net basic effect the contrary must occur. The amino acid must contain an extra amine functional group such as in the cases of lysine, arginine, and histidine. There is an exception to this rule in the case of tryptophan. Even though tryptophan has an amine group as part of a five member ring, the electron withdrawing effects of the two ring systems do not allow nitrogen to act as a base. Also, it should be noted that. Amino acids with an amide on the side chain do not produce basic solutions i.e. asparagine and glutamine. Although similar to an amine group, an amide has the carbon double bond oxygen which changes the property. Amides are neutral.

It is the combined effect of all the side chains which help to determine a proteins secondary shape and chemical reactivity. Those residues with acidic side chains will be more attracted to those with basic side chains so are expected to be in close proximity. It is likely that the charged residues will be present more frequently on the surface of the protein as charge (alongside shape and size) is another way in which a protein interacts with its corresponding ligand.

The Ramachandran Plot

With regards to the steric constraints of the peptide bond, Ramachandran *et al* demonstrated that there are 2 angles the polypeptide chain can rotate on either side of rigid peptide bond: Ø (phi) and Ψ (psi). Ø (phi) is the degree of rotation at the bond between the nitrogen and carbon atoms of the main chain whereas Ψ (psi) is the degree of rotation between the alpha carbon and carbonyl carbon atoms (Ramachandran *et al*. 1963). He developed a plot in order to visualize the dihedral angle ψ against φ of amino acid residues in a protein structure. The plot shows - in theory - which values or conformations of ψ and φ are possible for a residue in a protein. It also shows the empirical distribution of data points observed in a single structure. The shading on the plot represents the different regions described by Morris *et al.* (1992). The darkest areas correspond to the "core" regions and indicate the most favourable combinations of phi-psi values. Ideally, one would hope to have over 90% of the residues in these "core" regions as percentage of residues in the "core" regions is one of the better guides to stereo-chemical quality.

The different regions of the plot were taken from the observed phi-psi distribution for 121,870 residues from 463 known X-ray protein structures. The two most favoured regions are the "core" and "allowed" regions which correspond to 10° x 10° pixels having more than 100 and 8 residues in them, respectively. The "generous" regions were defined by Morris *et al.* (1992) by extending out by 20° (two pixels) all round the "allowed" regions. In fact, the authors found very few residues in these "generous" regions. Therefore they can probably be treated much like a "disallowed" region and any residues found in them should be investigated more closely. Excluding the amino acids proline and glycine, Ramachandrans plot shows that there is a limited amount of allowed confirmations that the peptide backbone can take. The planar and asymmetric nature of the backbone cause steric hindrance and limit the value that each residue can assume. It is the combined phi-psi values of each residue in the sequence which dictate the secondary confirmation of the back bone **(Figure 4)**.



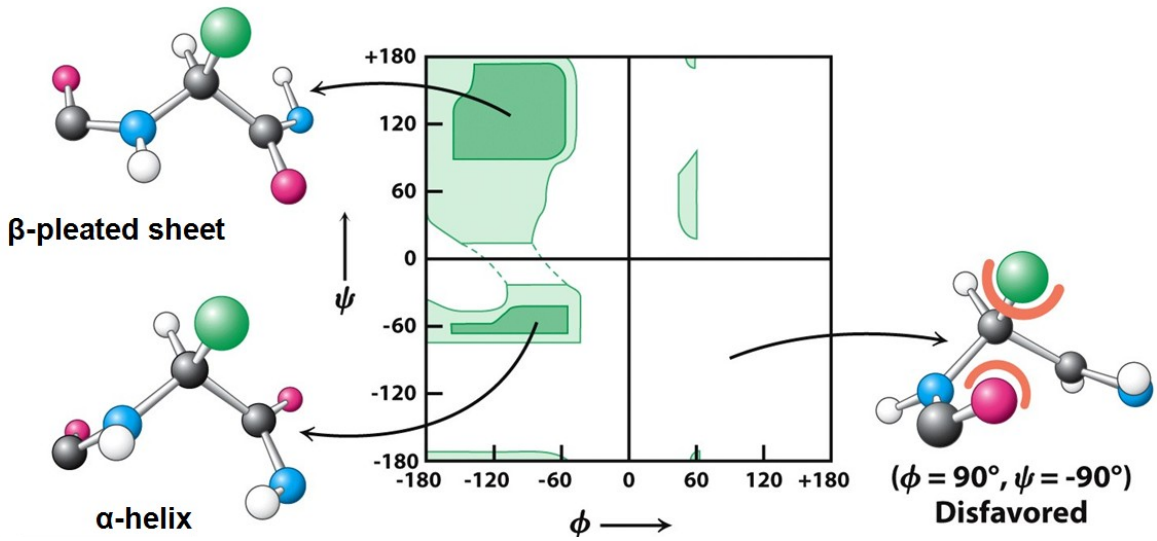


Figure 4: (Above) Image of a peptide chain showing the 2 angles that the chain can rotate on either side of the rigid peptide bond. The phi and psi angles are labelled. The green lines show the plane that does not rotate due to the double bond qualities of the peptide bond. (Below) An example of a Ramachandran plot of a peptide. Unlike a topographical map, energy instead of altitude is shown within the contours. Surrounding phi=0, psi=0 there is a high energy "plateau" which drops into valleys of stability with minima for alpha helices and beta sheet. In large proteins, the large majority of non-glycine residue possesses phi, psi combinations that reside in these valleys (Berg, Tymoczko, and Stryer 2012).

In reality there are only two major folding types possible in protein secondary structure. The first of these is called an alpha helix (α-helix) and the second a beta pleated sheet (β-sheet). Although a third formation - β-turn – is regularly seen, 60% of protein exists predominantly as helices and sheets. The remaining 40% experiences little if any periodicity and so acts as supportive elements (Lodish, Berk, and Kaiser 2007).

The α-Helix

Nuclear magnetic resonance (NMR), microscopy and x-ray crystallography have been used to analyse the secondary structures further. They have shown that the α-helix consists of 3.6 residues per turn and is stabilised through inter-strand hydrogen bonding parallel to the axis of the helix. Bonding occurs between the nitrogen atom in a peptide bond with the oxygen 4 residues down the chain. The distance between adjacent amino acid residues in the backbone is approximately 1.5Å – the length of a hydrogen bond – and due to the chiral nature of the amino acids, helices are also chiral. An α-helix is specifically right handed. The variable side chains in a helix point outwards and thus allow for compactibility of the structure. Also, the side chains are responsible for dictating the hydrophobic/hydrophilic nature of the helix (Bolsover et al. 2003).

There is an exception to this in the case of the amino acid proline. Proline, in general, cannot take part in a helix although it is sometimes present as one of the first four residues. Its looped side chain prevents the α-nitrogen from forming hydrogen bonds and a kink is produced instead (Woolfson and Williams 1990).

The β-Pleated sheet

In contrast to the relatively compact helix, the β-pleated sheet is almost fully extended. It is referred to as 'pleated' as geometry prevents it from being completely flat. The sheet is composed of adjacent but separate strands, 5-8 residues long, bound together in the form of a sheet. The sheet is stabilised by hydrogen bonding between the amine (N-H) and carboxyl (C=O) groups of adjacent strands with the side chains protruding above and below the sheet alternatively. The axial distance between adjacent amino acids is 3Å and the β-sheets can be described as either parallel or anti-parallel depending on the direction of each strand. Those with strands running in the same direction as each other are referred to as parallel where as those running in opposite directions are anti-parallel (Sipe, 2006).

The anti-parallel form is preferred over the two. This is due to that fact that when in this form, the alternate directions of the β-strands allow the N-terminus of one strand to be directly adjacent to the C-terminus of the next. This produces the strongest inter-strand stability as the hydrogen bond between the carbonyls and amines are planar. In the case of parallel strands, all the N-termini of successive strands are oriented in the same direction. This is unfavourable as it introduces non-planarity in the inter-strand hydrogen bonding pattern (Gailer and Feigel 1997) **(Figure 5) (Figure 6)**.

β-Turns

Beta turns are formed of 4 residues and are located on the surface of proteins. They are stabilised by hydrogen bonding between the end residues and their structure provides compatibility of the protein structure. They are formed predominantly by glycine and proline residues. Each of these lack a large side chain and thus allow close proximity of the beta turn residues and the protein as a whole (Lodish, Berk, and Kaiser 2007).

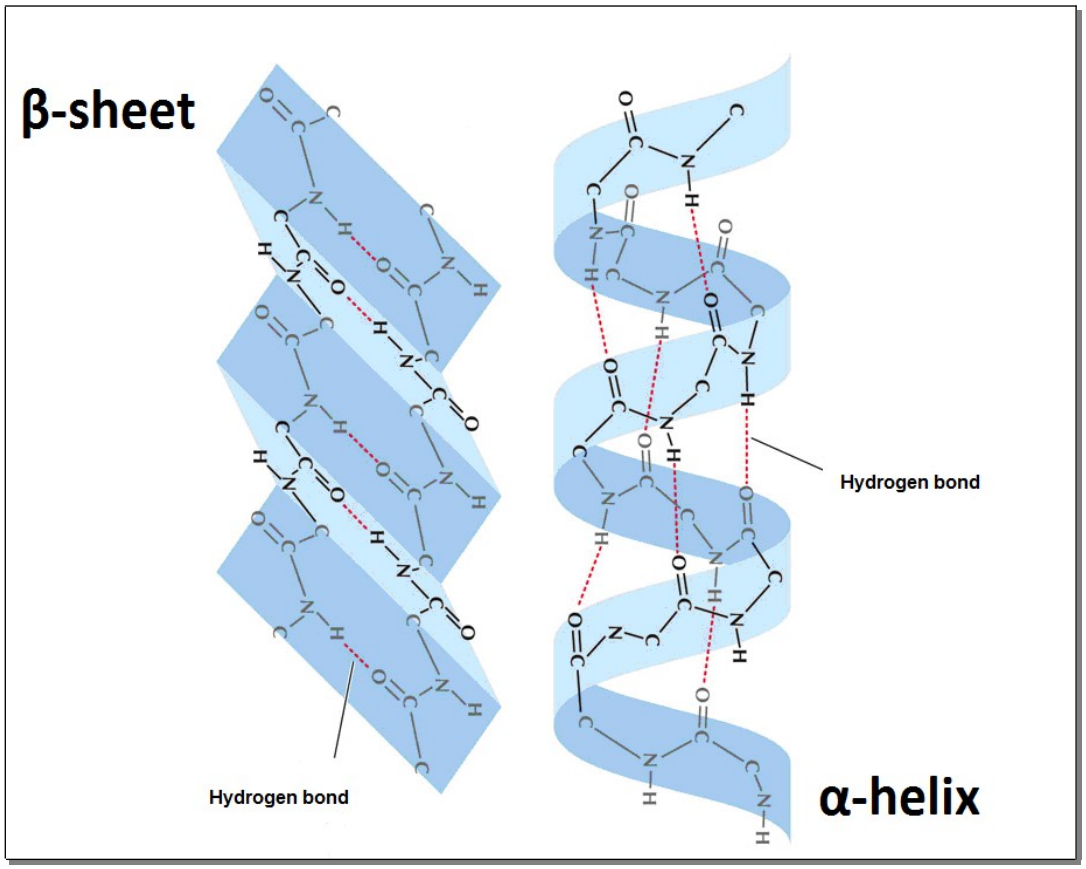


Figure 5: A detailed look at the α-helix and β-pleated sheet. The α-helix consists of 3.6 residues per turn. This allows the nitrogen atom in each peptide bond to form a hydrogen bond with the oxygen 4 residues down the chain. The side chains point outwards in a helix allowing compactibility. The distance between adjacent amino acid residues in the backbone is 1.5Å. An α-helix is right handed due to the chiral nature of the amino acids. The β-pleated sheet is composed of adjacent but separate strands 5-8 residues long bound together in the form of a sheet. The sheet is stabilised by hydrogen bonding between the amine (N-H) and carboxyl (C=O) groups of adjacent strands with the side chains protruding above and below the sheet alternatively. The axial distance between adjacent amino acids of 3Å. Strands running in the same direction as each other are referred to as parallel where as those running in opposite directions are anti-parallel (Madej et al. 2014) (Cooper and Hausman 2009).

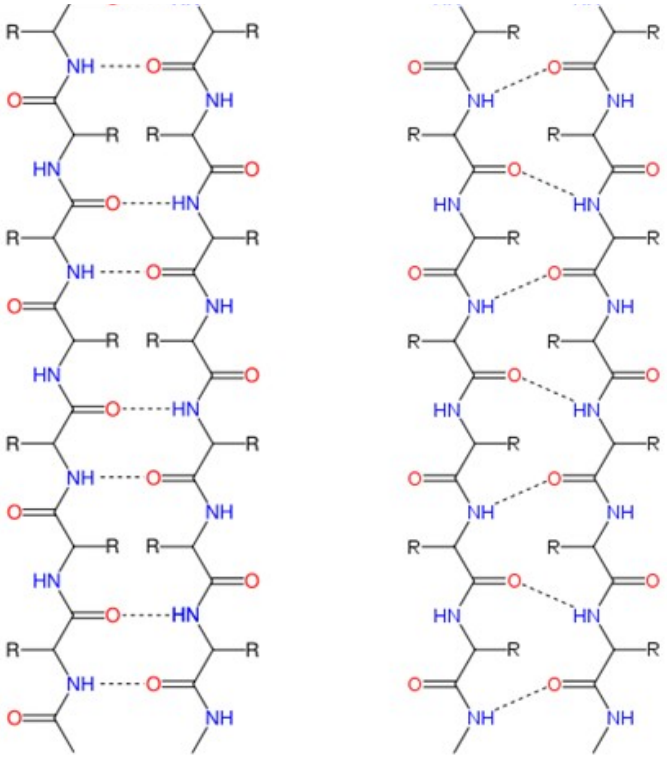


Figure 6: Diagrams representing the two types of β-sheet. Anti-parallel sheets are represented on the left. They occur when the strands run in opposite directions. Parallel strands are shown on the right and they occur when the strands run in the same direction. Both sheets are almost fully extended and stabilised by hydrogen bonding between the amine (N-H) and carboxyl (C=O) groups of adjacent strands. The anti-parallel form is preferred as when in this form, the N-terminus of one strand is directly adjacent to the C-terminus of the next. Hydrogen binding is planar and strong. In the case of parallel strands all the N-termini of successive strands are oriented in the same direction. This is less stable as inter-strand bonding is non-planar. The side chains of each protrude above and below the sheet alternatively and the axial distance between adjacent amino acids is 3Å.

Tertiary Structure

Generally, a protein folds into a mixture of α-helices and β-sheets although is it possible to be constructed exclusively of one or the other. The term tertiary structure refers to the overall shape that a single protein molecule will make; *i.e.* the spatial relationship of the secondary structures to one another.

Unlike secondary structures, which are stabilized by strong hydrogen bonds, tertiary structures are stabilized by non-local interactions. These interactions are influenced most commonly by the formation of a hydrophobic core, but also through salt bridges, hydrogen bonds, disulfide bonds, and even post-translational modifications. Although considered 'weak', compared to hydrogen bonding tertiary structure bonds are much higher in number so overall produce a stable structure (Bolsover et al. 2003).

The strongest of these tertiary influences would be the formation of the hydrophobic core. Protein folding is critically dependent on the effect of solvent (water) on non covalent interactions. Side chains which react strongly with water are called hydrophilic residues and those which do not are called hydrophobic. Hydrophobic molecules force water around them to become more organised in the form of cages. This is undesirable as is disrupts the hydrogen binding in bulk water. In protein folding, cage formation is minimised by the clustering together of such molecules into a hydrophobic core in something named the 'hydrophobic effect'.

Thermodynamic studies have shown that a polypeptide with hydrophobic and hydrophilic residues will spontaneously adopt a configuration where hydrophobic residues are not exposed to water either though a lipid bi-layer or through central clustering of those residues. *i.e.* Polypeptides prefer folding in the context of the solvent rather than be fully extended and having maximum interaction with the solvent (Sharp, 1991). If positively and negatively charged residues are buried in the hydrophobic interior core they will attract each other and the force between them will be stronger than if they had interacted with water. Such an electrostatic interaction is called a salt bridge.

Protein Folding Thermodynamics

Although the steric restrictions of the peptide bond and the hydrophobic effect are both essential in creating a compact and stable structure, neither is actually responsible for initiating the protein folding pathway. What drives the direction of the reaction specifically from unfolded to folded? How does the protein know when it has obtained its final, native state? Both these questions can be answered using thermodynamics.

All natural processes are subject to two laws of thermodynamics.

1 – Energy is conserved (it cannot be made or destroyed)
2 – Entropy is always increasing.

The laws of thermodynamics govern the energy and entropy differences in processes in which a system goes from one stable form to another. The first law does not distinguish between the *direction* of the reaction but the second law does. The reverse of a spontaneous process is not a natural process. This can be looked at in terms of heat. If a hot object is put in contact with a cold object, heat will spontaneously travel from the hot to the colder one (law 2) and the heat absorbed by the cold one will be equal to amount given (law 1).

The second law tells us which processes are spontaneous and thus tells us about equilibrium. Equilibrium is described as being the state of a chemical reaction in which its forward and reverse reactions occur at equal rates. Overall this means that the concentration of the reactants and products does not change with time.

Equilibrium is governed by Gibbs free energy:

$$G = H-TS$$

H = enthalpy, T = temperature and S = entropy

A system held at constant temperature and pressure will be at equilibrium if the Gibbs free energy is at a minimum. At the minimum, any infinitesimal change in the state of the system will leave G unchanged, therefore $\Delta G = 0$ is the criterion for equilibrium at any constant temperature and pressure (Lesk 2010).

For a system to be in equilibrium it must be in a state that is not the initial state of a spontaneous reaction and a compromise of demands of energy and entropy must be made. For example, in terms of proteins, the equilibrium between native and denatured states of a protein at the melting point is a balance between the higher cohesive forces in the compact native state and the greater conformational freedom in the ensemble of denatured conformations.

The Levinthal Paradox

When in its native form a protein is in its most stable form as the Gibbs free energy is at its lowest point. However, protein structures are only stable and functional over a small range of environmental conditions. Outside this range the interactions destabilise, the structure is disrupted and the molecule denatures.

Denaturation can be caused by a number of factors. These include excessive temperatures, sequence mutations, fragmentation, pH changes and detergents. Once denatured, the protein usually remains in its new form even after the environmental conditions have been rectified. The exception to this is when using urea. As expected, upon exposure to 6-8M urea, non covalent interactions are disrupted and proteins lose their higher level of structure. The polypeptide chains adopt random, changing conformations and there is an increase in entropy. However, upon removal of the urea, either by dialysis or dilution, the protein refolds and regains its original structure and activity (Bolsover et al. 2003).

The ability of the protein to return to its native state indicates that the final structure is pre-determined by the sequence and that refolding after denaturation is not random. A small protein of only 100 residues would take 10^{50} years to attempt every structural conformation, not the seconds it takes in vivo. This can be described by the Levinthal paradox. This states that proteins cannot fold by random search in conformation space. Proteins must have effective ways of folding encoded in their sequence. A folding pathway must exist (Zwanzig, Szabo, and Bagchi 1992).

Unfortunately, it is extremely difficult to visualise this pathway as, in denaturation experiments, a two stated equilibrium is observed with no stable intermediates; *i.e*. upon the increase of temperature or denaturant, there is a sharp transition from the native to the denatured state. If no stable intermediates are observed it is impossible to map the pathway taken for protein folding. Is there a single route per protein or are there various?

The current view of the folding process is that of a funnel. The funnel is wide at the top to encompass the different denatured conformations (high energy) and narrow at the bottom the represent the single native form (low energy). You can get to the native from the denatured via many routes but these eventually coalesce at the narrow part of the funnel (Bryngelson et al. 1995). Although there is not necessarily a single direct pathway for protein folding, certain factors must influence the final structure. For example, interactions between regions nearby in the sequence should form more easily than those in the distance. This provides guidance for the secondary structures to form. Secondary structures are more stable than unfolded structures and once formed, the globular protein only needs to compact itself **(Figure 7)**.

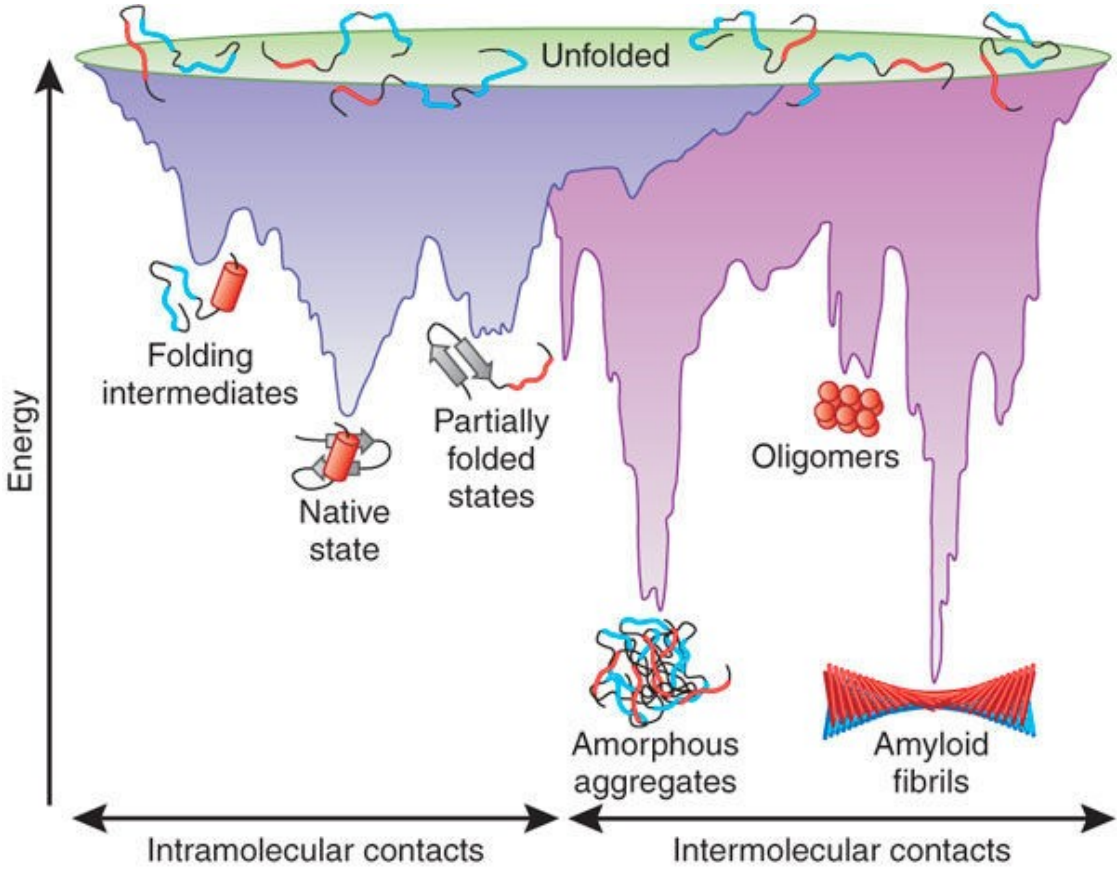


Figure 7: The current view of the folding process - a funnel. The blue surface shows the multitude of conformations 'funnelling' to the native state via intramolecular contacts. The funnel is wide at the top to encompass the different denatured conformations (high energy) and narrow at the bottom the represent the single native form (low energy). The purple area shows the conformations moving toward amorphous aggregates or amyloid fibrils via intermolecular contacts. Both parts of the energy surface overlap. Aggregate formation can occur from intermediates populated during de novo folding or by destabilization of the native state into partially folded states and is normally prevented by molecular chaperones. Cell-toxic oligomers may occur as off-pathway intermediates of amyloid fibril formation (Hartl and Hayer-Hartl 2009).

To go from a higher energy (denatured) to lower energy (native) the proteins sometimes have to overcome an energy barrier to enter the transition state. In a regular chemical reaction one would overcome this barrier by increasing the pressure, concentration or temperature of the solution. In doing so it increases the energy and rate at which particles collide and hence increases the likelihood of a reaction. Although effective, each of these methods are not always ideal for protein folding as fluctuations in temperature and pressure often cause secondary and tertiary structure disruption. In order to overcome the energy barrier without effecting protein fold a chaperone is typically used (Hartl and Hayer-Hartl 2009).

Chaperones

Chaperones are protein molecules which act as folding catalysts. They are a large group of unrelated proteins with a range of roles all related to protein folding. These include de novo folding of polypeptides, refolding of mis-folded proteins, solubilization of protein aggregates, degradation of proteins, translocation of proteins across membranes, assembly and disassembly of oligomeric complexes and the regulation of stability and activity of certain natively folded proteins. Some chaperones are non-specific and act with a variety of polypeptide chains where as others are restricted to specific targets. They often use ATP hydrolysis to assist in intermediate folding but they do not interact with native proteins and they do not form part of the final folded structure (Gething and Sambrook 1992).

Chaperones are evolutionarily conserved across all organisms highlighting their importance in *in vivo* protein folding. There are two types of chaperones: molecular chaperones, which bind and stabilise unfolded proteins preventing aggregation, and chaperonins which form a small folding chamber into which an unfolded protein is presented to the appropriate environment for folding.

Molecular Chaperones

The heat shock proteins are examples of molecular chaperones. They range in size between 10-100kDa and are named according to their size. When bound to ATP the heat shock proteins assume an open form which exposes central hydrophobic regions. The hydrophobic regions of an unfolded protein bind to this and cause the hydrolysis of the ATP to ADP. Once folded, the protein is released and the chaperone reverts back to its closed position.

Take for example the heat shock protein Hsp70. As the name suggests, HSP70 is 70kDa in size. It is composed of three domains: the N-terminal ATPase domain, the substrate binding domain and the C-terminal domain. The n-terminal ATPase domain is the location for ATP binding. The energy molecule lies at the bottom of a deep cleft held between two lobes. It is the exchange of ATP and ADP which is responsible for the conformational changes of the other two domains into the open and closed positions (Mayer 2010).

In Hsp70 the substrate binding domain is composed of an α-helix subdomain 10kDa in size and a β-sheet subdomain 25kDa in size (Zhu et al. 1996). The β-sheet domain consists of two twisted sheets arranged into a barrel with its connecting loops protruding upwards. It is in this barrel that the protein intermediate is contained. The intermediate is held in place within the substrate binding domain via a groove with affinity for hydrophobic residues. The groove is long enough to interact with peptides up to seven residues in length.

The final c-terminal domain acts as a lid for the substrate binding domain. It is composed of five α-helices. Helices A and B pack against two sides of the β-sheet sub-domain, stabilizing the inner loops $L_{1,2}$ and $L_{4,5}$. In addition, helix B forms a salt bridge and several hydrogen bonds to the outer Loops L3,4 and L5,6, the so-called latch, thereby closing the substrate-binding pocket like a lid. Helices C through E, together with the distal part of helix B, form a second

hydrophobic core. The function of this second hydrophobic core has yet to be defined however it is believed to aid in the stabilisation of helix B, since removal of helices C through E leads to unfolding of the distal part of helix B (Jiang et al. 2005).

When a Hsp70 protein is ATP bound, the lid is open and peptides bind and release relatively rapidly. When Hsp70 proteins are ADP bound, the lid is closed and peptides are tightly bound to the substrate binding domain.

Chaperonins

In contrast to molecular chaperones, chaperonins are huge cylindrical structures formed of two rings of oligomers. GroEL is an example found in *e.coli*. It is a homo-oligomer formed of 7 identical subunits per ring, 57kDa in size, packed back to back. An unfolded protein of 60kDa or smaller enters one of the rings and binds to the inner wall via hydrophobic interactions with several of the subunits. From then on the ring containing the substrate is referred to as the “cis” ring. When ATP binds to the cis ring it allows the co-chaperone GroES to bind and enclose the substrate within. Formation of GroEL-GroES requires a large conformational change and causes the interior surface of cis to go from hydrophobic to hydrophilic giving the unfolded protein a chance to fold. After approx 20 seconds ATP is hydrolysed and GroES and ADP are released leaving the barrel to take a more open conformation (Horwich et al. 2007). The protein is released whether it has folded or not. At which point it can either re-enter the same chaperonin or enter another to try again. The cis and “trans” ring act allosteric to one another. Substrate enclosure in one ring leads to substrate release in the other (Yifrach and Horovitz 1995) **(Figure 8).**

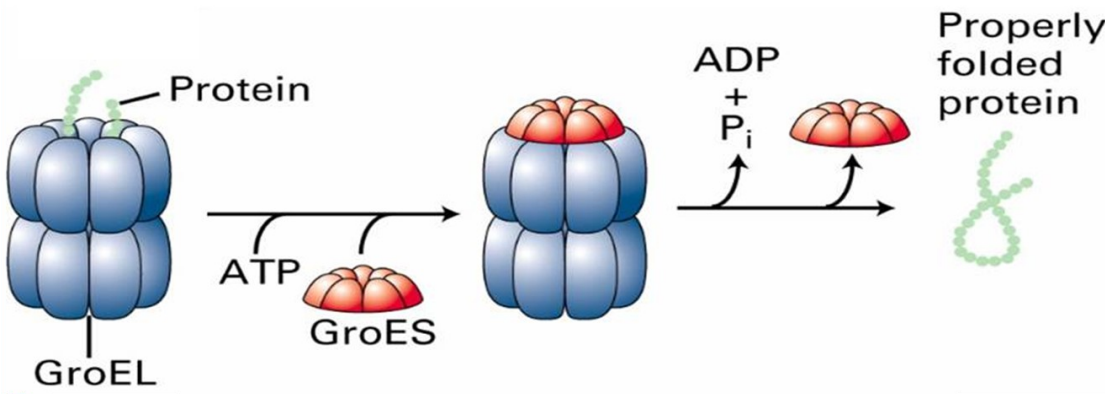


Figure 8: A diagram showing assisted protein folding by the chaperonin GroEL. GroEL is formed of 2 rings each containing 7 subunits packed back to back. As an unfolded protein enters a barrel, an ATP molecule and the assistant chaperone GroES enclose the structure and induce a large conformational change. Formation of GroEL- GroES causes the interior surface to go from hydrophobic to hydrophilic giving the unfolded protein a chance to fold. After approx 20 seconds the ATP is hydrolysed releasing a phosphate molecule and the GroES cap. This conforms the ring into a more open form, causing it to release the protein whether it has folded or not. At this point it can either re-enter the same chaperonin or enter another to try again (Lin, Madan, and Rye 2008)

An Introduction to Amyloid

When one takes into account the vast array of jobs performed by proteins; receiving cell signals, regulation of waste and maintaining structure of tissues and organs, it becomes quite clear why the body makes great efforts to ensure a protein reaches its native state in a fast and reliable manner. Proteins are responsible for almost every task of cellular life but they can only carry out these tasks when in their native state. This is because the native state is the only structural formation in which the proteins compliment their three dimensional ligands in shape and charge. For this reason, any disruption of the three-dimensional shape can lead to disruption in cell regulation. Some of these disruptions are manageable. Drugs can often be used to control the physiological imbalance and reduce any symptoms however this is not always the case. Sometimes the effects of protein denaturation can be more destructive and have a more consequential outcome; *i.e.* a disease can develop.

This thesis focuses on a specific group of protein related diseases; amyloid.

Amyloidosis is the name given to a group of specific protein mis-folding diseases. It occurs when proteins unfold from their native state and fold into a highly regular cross β-sheet aggregate structure known as an amyloid fibre. This cross β-sheet fibre structure is present in all amyloid regardless of the protein involved. Amyloid fibres are insoluble, rigid and thermodynamically stable compared to their native counterparts and once formed, amyloid fibres associate with extracellular matrix proteins and proteoglycans to form structures referred to as amyloid deposits. These deposits invade tissues and organs disrupting tissue structure and function (Sipe 2006) **(Figure 9)**.

Figure 9: Photograph of a kidney tarnished with sticky, yellow amyloid deposits (Gertz and Rajkumar 2013).

The mechanism behind amyloid tissue damage is not fully understood, however, it is generally believed to be a mechanical or structural problem rather than a cytotoxic one. Cytotoxic protofilaments *are* present *in vivo*, however, it is the structurally disruptive deposits which appear to have a greater contribution to disease, *e.g.* the deposition of β2M amyloid in the glomerulus stiffens this vessel, preventing blood from being filtered in the kidneys (M. B. Pepys 2001). Be that as it may, more recent evidence has suggested that at least in the case of Alzheimer's disease, the presence or absence of amyloid plaque is insufficient to fully account for the damaging role of elevated Aβ protein in amyloid-beta derived diffusible ligands (ADDLs). The soluble oligomers of Aβ - the ADDLs - accumulate and cause functional defects prior to neuronal cell death or plaque deposition (Catalano et al. 2006).

Various conditions are known to induce amyloid fibre growth. These include fragmentation of the protein, mutations and inheritance of a degenerative protein variant. Further causes include deamination of asparagines and glutamine into their aspartic and glutamic acid counterparts, oxygenation of the backbone carboxyl and amine groups and unregulated translational modifications such as methylation and hyperphosphorylation (Goedert 1993). All of which result in an unstable structure, prone to mis-folding, being formed.

Even age can be another inducing factor for protein instability as with the development of senile systemic amyloidosis (SSA). In this particular case, the mutated Transthyretin (TTR) protein only becomes amyloidogenic after years of accumulation (Westermark et al. 1990).

The History of Amyloid

Amyloid structures were originally discovered in 1854 by Virchow and Meckel through their histopathological studies on human patients. They originally mistook it for starch and used iodine staining as an analytical technique to debate as to whether amyloid had a cellulose or cholesterin nature (Virchow

1854). However, Friedreich and Kekule's subsequent chemical analysis in 1859 revealed that amyloid deposits were actually of protein origin and this was verified 50 years later when Hanssen successfully digested amyloid with pepsin – a protein digestion enzyme. In 1922, Bennhold introduced the Congo red staining method of detecting amyloid. The associated green birefringence suggested that amyloid was not an amorphous aggregate as originally thought, but instead an organised pattern of protein (Sipe 2006). Cohen and Calkins confirmed the fine ultrastructure using electron microscopy (EM) in 1959 but it was Glenner in 1975 who confirmed that the structure was cross ß-sheets using x-ray diffraction. This cross ß-sheet motif is present in all amyloid deposits and is identical in structure regardless of the protein involved (G G Glenner 1975).

Amyloid Diseases

Although the structure of amyloid fibre is always cross β-sheet in nature, the location of the amyloid deposit varies depending on the protein involved. As a result of this, it is not always possible to predict which organ will be affected making diagnosis a difficult task. If multiple organs are affected it is referred to as systemic amyloidosis. This is the most common form of amyloidosis and it can occur in any organ with the exception of the intra-cerebral area (Mark B. Pepys 2006). Various examples of systemic amyloidosis are serum amyloid A amyloidosis (AA) amyloid light chain amyloidosis (AL), and TTR amyloidosis (ATTR). AA amyloidosis is the most common type worldwide however systemic AL amyloidosis is the most common type in developing countries accounting for approximately 55% of the patients in the UK alone (Real de Asúa et al. 2014).

In the case of AL amyloidosis, a plasma cell dyscrasia leads to the over production of immunoglobulin chains (light or heavy) with the propensity to form amyloid. Although AL amyloidosis it is one of the most common forms of systemic amyloidosis, its wide range of symptoms can make it difficult to

diagnose. AL has the largest spectrum of organ involvement of all the amyloid diseases but most typically affects the kidneys, heart and peripheral nervous system (PNS). It causes symptoms such as fluid retention, shortness of breath, heart failure and palpitations; symptoms which are more commonly associated with coronary heart disease, bronchitis and pneumonia. Other symptoms can include stroke, diminished organ function and weight loss. It is diagnosed through the detection of high concentrations of light chains in the blood and urine and is treated via chemotherapy and bone marrow transplantation (Gertz and Rajkumar 2013). In contrast to this, Aβ2M typically affects the kidneys and the musculoskeletal system whereas ATTR generally affects the heart and liver.

In the rarer cases where amyloidosis is restricted to a single organ it is referred to as localised amyloidosis. This form of amyloid is usually secondary to a non-infectious chronic inflammatory disease and is caused by localised excess of hormone protein. An example of this is during insulin amyloidosis. Here, large concentrations of insulin are noted at the injection site of diabetics. The deposits are seen to form a hard nodule at the injection site and are known to release unpredictable amounts of insulin at any given time. Cases of severe hypoglycaemia have been reported due to the irregular absorption from the affected site and weight gain is commonly observed. In these cases treatment involves removal of the hard nodule through surgery and avoidance of injections at the amyloidosis site (Sipe 2006).

Proteins Associated with Amyloidosis

At present, a total of 36 proteins have been associated with amyloid disease and cross β-sheet folding *in vivo* according to the International Committee on Nomenclature of the Amyloid Proteins (Sipe et al. 2016). Nevertheless, Rochet *et al* have shown that non amyloidogenic proteins can also form fibres when subject to low pH and high temperatures *in vitro* (Rochet, Conway, and Lansbury 2000). Based on these findings Dobson in 2002 suggested that

fibrillisation is a component of all protein molecules and not just those involved in disease. Dobson believes that fibrillisation can be triggered in any protein if they were to become exposed to the correct conditions **(Table 1)**.

The human body hosts more than 50,000 proteins, each of which has a unique amino acid sequence. It is difficult to understand how proteins of unique tertiary structures and sequences would be able to form a universal structure. A protein sequence of 10 residues alone already processes 10^{20} sequence possibilities not to mention proteins with hundreds of residues. Even proteins which are composed of mostly β-sheets such as TTR and β2-microglobulin would still require major conformational changes to convert to the cross β-sheet structure never mind proteins composed primarily of α-helices.

The existence of the cross ß-fold amongst such a diverse array of proteins demonstrates its superior thermodynamic stability compared to wild type proteins and explains its resistance during treatment. However, when researching amyloidosis one will notice that there are certain points which are not explained in the literature. Scientists are uncertain if all proteins truly have the potential to form amyloid. Also, it is unclear as to why amyloid forms at all when the native state of a protein is so stable.

The fact that proteins do not all spontaneously form fibres, especially *in* vivo, could be explained by a number of points: The first being the presence of chaperones *in vivo*. These assist in the folding of peptide chains into functional native states. The second reason is that physiological conditions do not necessarily induce the initial unfold required of proteins to make the cross β-sheet fold. One must remember that the native state of a protein is a highly stable structure which is not easy to disrupt. A pH of 7.4 and a temperature of 37°C are maintained in a body. These conditions are not likely to induce protein mis-folding without the additional factors of sequence mutations, and fragmentation.

Fibril Protein	*Precursor Protein*	*Systemic / Localized*	*Acquired / Hereditary*	*Target Organs*
AL	Immunoglobulin light chain	S, L	A, H	All organs, usually except CNS
AH	Immunoglobulin heavy chain	S, L	A	All organs except CNS
AA	(Apo) Serum amyloid A	S	A	All organs except CNS
ATTR	Transthyretin, wild type	S	A	Heart mainly in males, ligaments, tenosynovium
	Transthyretin, variants	S	H	PNS, ANS, heart, eye, leptomeninges
Aβ2M	β2-Microglobulin, wild type	S	A	Musculoskeletal system
	β2-Microglobulin, variant	S	H	ANS
AApoAI	Apolipoprotein A I, variants	S	H	Heart, liver, kidney, PNS, testis, larynx (C-terminal variants), skin (C-terminal variants)
AApoAII	Apolipoprotein A II, variants	S	H	Kidney
AApoAIV	Apolipoprotein A IV, wild type	S	A	Kidney medulla and systemic
AApoCII	Apolipoprotein C II, variants	S	H	Kidney
AApoCIII	Apolipoprotein C III, variants	S	H	Kidney
AGel	Gelsolin, variants	S	H	PNS, cornea
ALys	Lysozyme, variants	S	H	Kidney
ALECT2	Leukocyte chemotactic factor-2	S	A	Kidney, primarily
AFib	Fibrinogen α, variants	S	H	Kidney, primarily
ACys	Cystatin C, variants	S	H	PNS, skin
ABri	ABriPP, variants	S	H	CNS
ADan	ADanPP, variants	L	H	CNS
Aβ	Aβ protein precursor, wild type	L	A	CNS
	Aβ protein precursor, variant	L	H	CNS
AαSyn	α-Synuclein	L	A	CNS
ATau	Tau	L	A	CNS
APrP	Prion protein, wild type	L	A	CJD, fatal insomnia
	Prion protein variants	L	H	CJD, GSS syndrome, fatal insomnia
	Prion protein variant	S	H	PNS
ACal	(Pro)calcitonin	L	A	C-cell thyroid tumors
AIAPP	Islet amyloid polypeptide	L	A	Islets of Langerhans, insulinomas
AANF	Atrial natriuretic factor	L	A	Cardiac atria
APro	Prolactin	L	A	Pituitary prolactinomas, aging pituitary
AIns	Insulin	L	A	Iatrogenic, local injection
ASPC	Lung surfactant protein	L	A	Lung
AGal7	Galectin 7	L	A	Skin
ACor	Corneodesmosin	L	A	Cornified epithelia, hair follicles
AMed	Lactadherin	L	A	Senile aortic, media
AKer	Kerato-epithelin	L	A	Cornea, hereditary
ALac	Lactoferrin	L	A	Cornea
AOAAP	Odontogenic ameloblast-associated protein	L	A	Odontogenic tumors
ASem1	Semenogelin 1	L	A	Vesicula seminalis
AEnf	Enfurvitide	L	A	Iatrogenic

Table 1: A list of the 36 proteins involved in human amyloidosis (Sipe et al. 2016).

Fibre Formation

As explained previously, all amyloid fibres share this common cross ß-sheet motif regardless of the protein involved, however the precise mechanism of cross β-sheet folding is not fully understood. Findings from Booth *et a*l. in 1997 indicate that cross β-sheet formation is achieved through an initial unfolding of the protein, either fully or partially, from its native state into an aggregation prone intermediate. Booth discovered this whilst experimenting with two thermically unstable variants of lysozyme. He found that they - unlike their wild type counterpart - were able to form amyloid fibres from partially folded intermediates. This was later confirmed by Dumoulin *et a*l. in 2005. Of course, in cases where the protein is natively unfolded *i.e.* amyloid β (Aβ), the initial unfolding witnessed with lysozyme is inapplicable. In that particular case, one would assume that cross β-sheet folding was due to an increased concentration and proximity of amyloid prone isomers **(Figure 10)**.

Fibre growth is very similar to crystal growth in crystallisation; it is a nucleation dependent form of oligomerisation. This type of growth creates a sigmoidal curve which can be divided into three stages: nucleation, growth and stationary.

Nucleation

The nucleation stage involves a protein solution exceeding a critical concentration of amyloid prone intermediates in order to form a stable nucleus. The formation of amyloid nuclei is responsible for the lag phase typically seen in fibre growth as it is the rate limiting step of the process. Once formed, the remaining intermediates quickly assemble into nuclei and are elongated to form single stranded protofilaments (Joseph T. Jarrett 1993). Until recently it was believed that protofilaments were elongated at both ends however, Heldt *et. al* have provided evidence to the contrary. Using Alexa Fluor 568 labelled insulin protofilaments, they found that the majority of protofilaments propagated along one end of the nuclear seed although some

cases were seen at both. This indicates that, at least in the case of insulin, protofilaments are asymmetric (Heldt, Zhang, and Belfort 2011).

It should be noted that the lag phase can be avoided through the addition of pre-formed fibre, *a.k.a.* “seeding”. This eliminates the nucleation step completely as the cross β-sheet nuclei are already present and act as a catalyst for subsequent fibre growth (Cohen et al. 2012). Seeding is normally performed between variants of the same protein however there have been cases of seeding between different proteins where significant tertiary structural similarity with the host protein is observed (Ma and Nussinov 2012). Both seeded and unseeded fibre growth can occur in fibrillisation, however, seeded growth is hyperbolic and is restricted to *in vitro* fibre assembly. Unseeded growth is more commonly seen and can occur both *in vivo* and *in vitro.* The growth phase of fibres occurs at an exponential rate until a plateau is reached and soluble intermediate concentration is diminished (Sipe 2006).

Growth

A minimum of two protofilaments will arrange themselves around a hollow channel to make a rope-like structure approximately 10nm in diameter and of indefinite length. This rope-like structure is referred to as an amyloid fibre (Mark B. Pepys 2006). It is not known for certain how many protofilaments are required to make amyloid fibres, however, studies have shown cases of fibres ranging from two to six protofilaments (Serpell et al. 2000). One would assume that the number is dependent on the protein involved. For example, if one believes the theory that fibres are similar (if not identical) in diameter, then aggregates of smaller proteins must require more protofilaments to form a fibre of similar size to that of larger, bulkier proteins. Regardless of this, all fibre organisation requires hydrogen bonding and thus is dependent on pH and ion composition of the surrounding environment (Fraser et al. 1991) **(Figure 11)**.

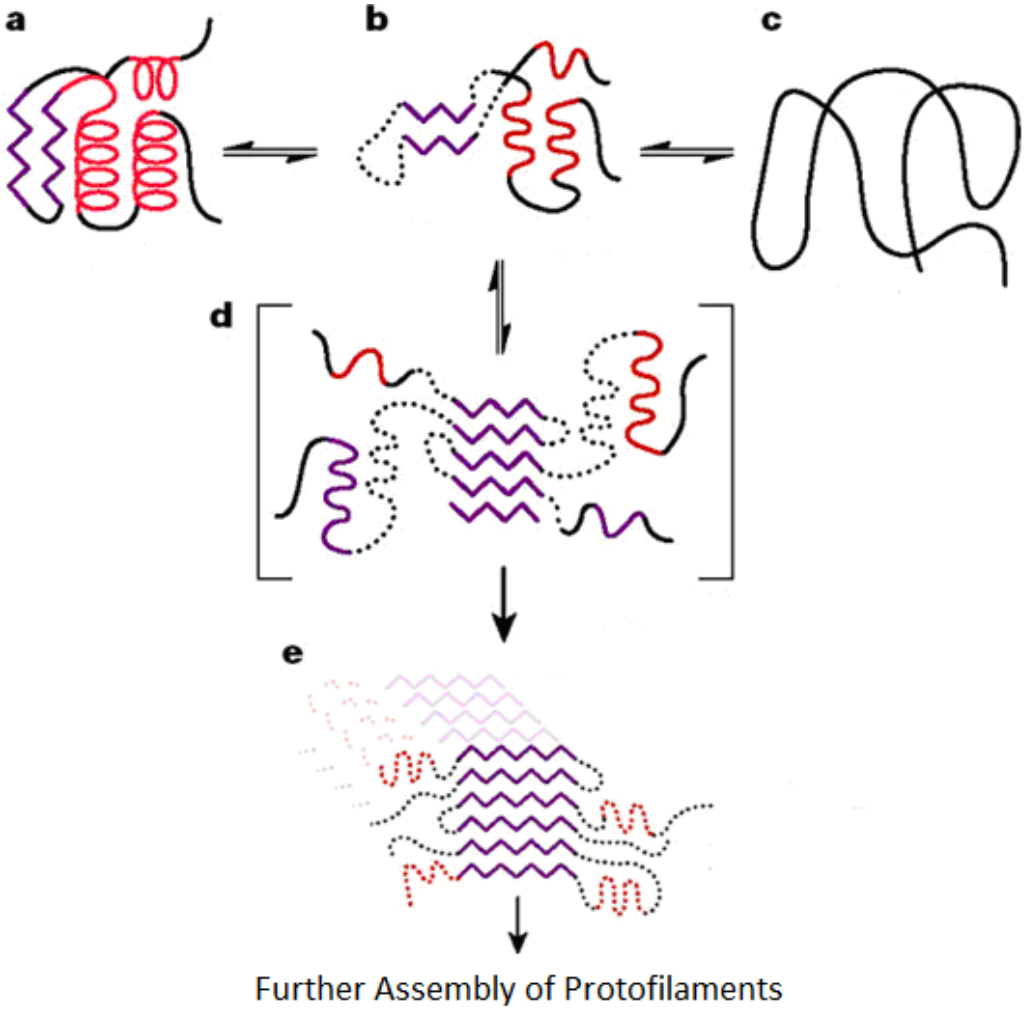


Figure 10: Proposed formation of cross β-sheet fibres from native protein state. Red represents α-helices, purple represents β-sheets and the dots represent disorganised regions (Selkoe 2003). Details of each step are listed below:

<u>Protofilament assembly:</u>

a – Protein in its native state

b – Partially unfolded intermediate. When in this state the protein may follow one of three steps: (a) Refold back into its native state through aid of protein chaperones. (c) Form amorphous aggregate with other unfolded protein monomers. (d) Spontaneously fold into a cross β-sheet nuclei.

c – Protein as an amorphous aggregate

d – Cross β-sheet Nucleus. Once formed, other unfolded intermediates are able to join at each end to create a long protofilament

e – Protofilament made of stacked β- sheets

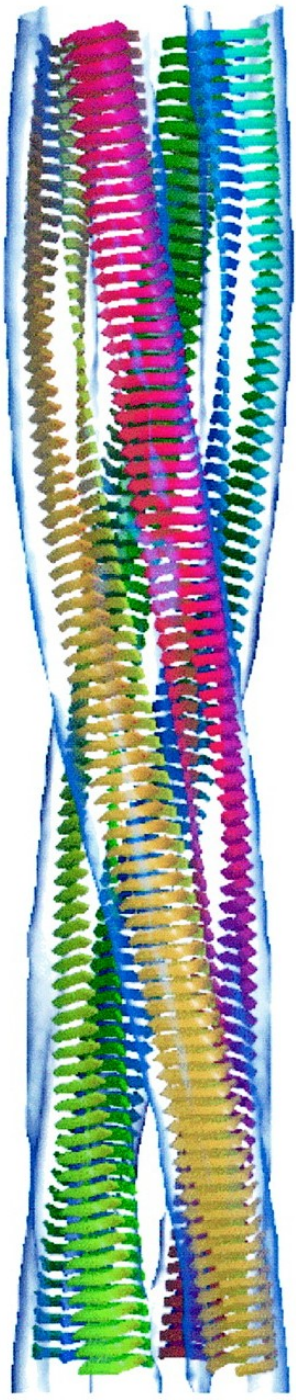


Figure 11: Model of an Insulin amyloid fibre composed of four protofilament strands. The strands are arranged around a hollow channel to make a rope-like structure approximately 10nm in diameter and of indefinite length (Jiménez et al. 2002).

Stationary

Once fully formed, the rate of growth decreases until a steady plateau is reached. In this so called stationary phase, the lack of visible 'growth' can cause one to believe that fibres exist as inactive structures. However, an interesting behavioural trait has been observed. Amyloid deposits are known to regress if circulating fibre precursors are depleted *i.e.* after transplantation or immunization (Holmgren et al. 1993) (Schenk et al. 1999) This shows that fully formed fibres are not inactive structures but are, in fact, dynamic. Many believe that the amyloid fibres are constantly disassembling and reforming themselves and that the rate at which fibres degrade and reform must therefore be equal. Only when the rate of growth falls below the rate of degradation will the fibres begin to break down. As of yet, there has been no clarity as to when the degradation process starts or to the duration of this stage. Due to the extreme stability of amyloid fibres, they are resistant to such an event. Any degradation would likely fall far from the time constraints of the sigmoidal curve; possibly years after growth.

Fibre Structure

When subject to x-ray analysis, the cross β-sheet fibre produces a smeared diffraction pattern instead of the spots typically seen in this type of analysis. This is due to the irregularity of the R-side chains. When analysed further, the results illustrate the cross beta sheet as a stacked structure containing multiple sheets of parallel β-strands. Instances of anti-parallel stacks have been noted in literature however this is quite rare. The distance between each sheet is 10.7Å and the distance between each strand in the sheet is 4.8Å - the equivalent of 3 sets of hydrogen bonding (Serpell 2000) **(Figure 12)**.

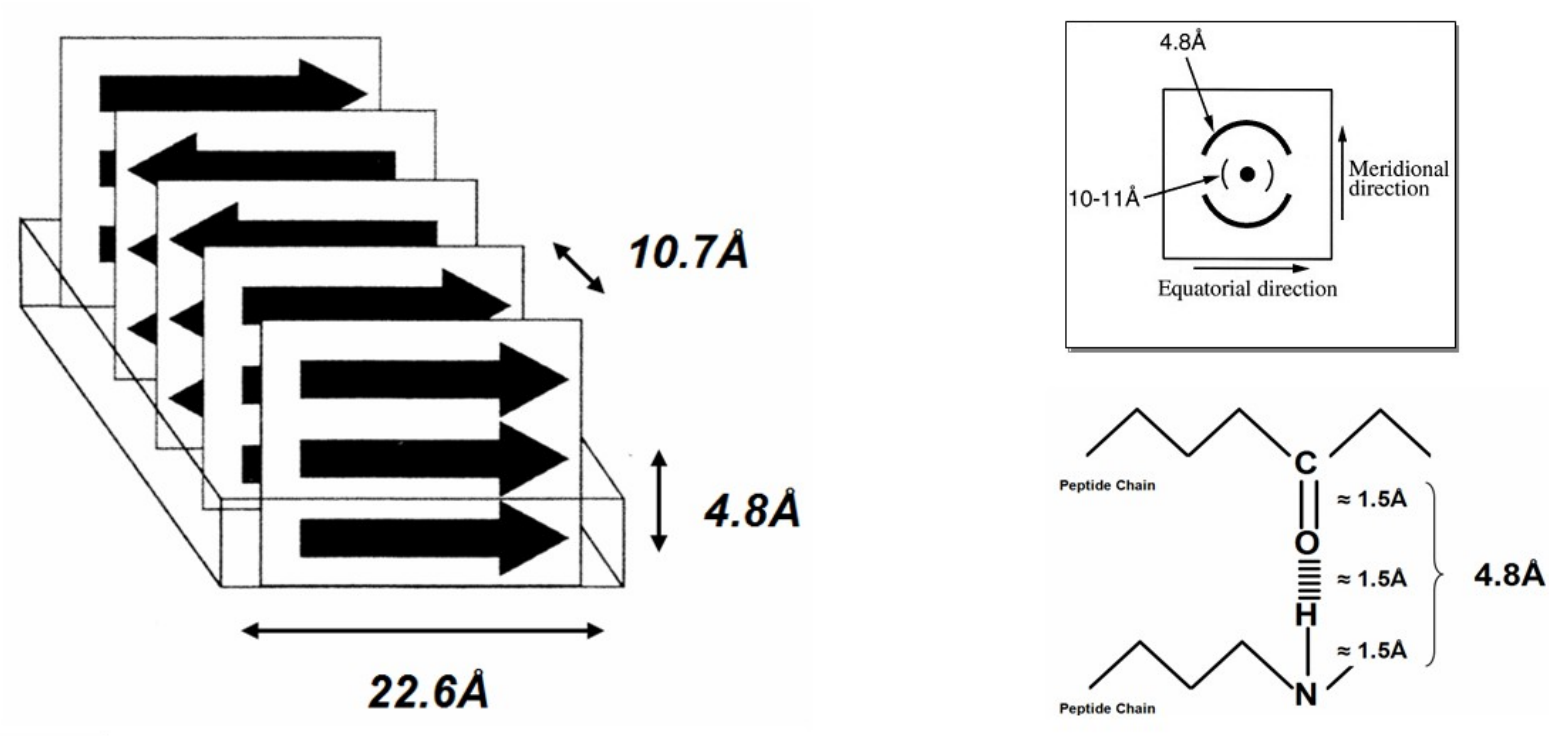


Figure 12: Left – A proposal for the arrangement of cross β-sheet fibres. They are composed of stacked β-sheets containing parallel β-strands (arrows). The distance between each strand in the sheet is 4.8Å - the equivalent of 3 sets of bonding (Right, bottom) - and the distance between each sheet is 10.7Å. This arrangement is believed to be identical regardless of the protein involved (Serpell 2000). There have been instances of anti-parallel stacks however this is rare as anti-parallel binding is less stable than parallel. Right (Top, boxed) – A typical x-ray diffraction pattern of parallel amyloid fibres. Note that the diffraction pattern appears as smears rather than the spots typically seen in this type of analysis due to the irregularity of the R-side chains.

The choice of parallel stacks in fibres instead of anti-parallel stacks is a surprising result. In regular protein folding, the anti-parallel form is preferred over the two due to that fact that hydrogen bonding between the carbonyls and amines are planar. It is unusual that the reverse is seen in cross β-sheets. Cross β-sheets have proven to be more stable than native proteins so one would expect them to be constructed of the more stable of the two β-sheet formations. Alternatively, one could argue that although the parallel form is less stable, it may be better suited for “stacking”. Perhaps stacking is more compact in the parallel form as the R groups are not directly opposite each other. Also, protofilaments twisting may be easier if hydrogen bonding in the stacks is not as rigid.

Native State Stability

The native state of a protein is described as the conformation at which the Gibbs free energy of the system is at its lowest; *i.e.* the native conformation is determined by the totality of inter-atomic interactions and hence by the amino acid sequence itself (Anfinsen 1973). With that said, cross β-sheets have shown to be even more thermodynamically stable. They, unlike their native counterparts, have the ability to maintain structure in extreme temperatures and in the presence of denaturing solvents provides.

Unfortunately is impossible to directly compare the thermodynamic stability of cross β-sheets with their native state equivalents. This is due to the insolubility, heterogeneity and high degree of polymerisation of the amyloid fibrils. However, β-sheet rich oligomers *can* be compared to native proteins as they, unlike regular cross β-sheets, are soluble.

Multiple studies have been performed using prion protein (PrP). The different states of PrP were analysed and used as a reference in the energy free diagram. The results show that the native protein state does not have the lowest Gibbs free energy and that the cross β-sheet is far more stable.

However, this cannot be taken as fact. The oligomers may not be the true global energy minimum as they themselves may undergo additional transitions to form polymeric amyloid forms (Hartl and Hayer-Hartl 2009).

Considering its supremacy on stability one would wonder why the cross β-sheet form is not accessible during folding under native conditions. This has already been discussed in literature and agreed to be due to the fact that β-sheets have a slower folding rate compared to folding of the native form. There must be a large energy barrier separating the native from the beta sheet (Baskakov et al. 2001). Also, analysis of the kinetic traces indicates that the process of folding into the β-sheet represents a transition with an apparent reaction order of 3 or higher (Sokolowski et al. 2003). An order of reaction is defined as a number that relates the rate of a chemical reaction with the concentrations of the reacting substances. The high order of reaction presented above suggests that transition depends highly on the concentration of the transition state.

Identification of Amyloid Fibres

There are various ways in which to identify amyloid fibres, the most common being the use of dyes Congo red and Thioflavin T (ThT).

Despite their common use in the diagnosis of amyloid fibres in ex vivo, in vitro, and animal model studies, not much is known about their mechanisms of amyloid binding. A large number of studies have examined the binding mode, but a high-resolution characterization has so far not been possible because of the insolubility and often heterogeneous nature of the amyloid fibres. In spite of this poor understanding, various theories have been presented throughout the years.

Thioflavin T

Thioflavin T (ThT) is a cationic benzothiazole dye that shows enhanced fluorescence upon binding to amyloid at a wavelength of 420nm. It is often used to diagnose amyloid in tissue sections using fluorescence microscopy and to monitor extracted amyloid and *in vitro* amyloid fibre formation using fluorescence spectroscopy **(Figure 13)**.

ThT has a minimal effect on the fibrillation kinetics and is independent of the number or length distribution of fibres. When analysing amyloid fibre formed from specific protein under given conditions, the fluorescence intensity of ThT has been shown to be proportional to the concentration of fibres. However, extreme care should be taken when comparing fibre formation under different conditions since a number of factors may affect the fluorescence intensity. These include the specific protein forming the fibril, fibril morphology, ThT concentration, pH, and ionic strength (H. LeVine 1997) (Harry LeVine 1995).

The concept of using ThT as an indicator of amyloid was first introduced by Vassar and Culling in 1959. They demonstrated that the detection of amyloid in the kidneys with ThT after differentiation in acidic solutions was highly specific. This was later confirmed in 1967 by Saeed and Fine. Unfortunately they found that ThT was also able to bind to other connective tissues such as cartilage matrix and elastic fibres. Furthermore, it demonstrated enhanced fluorescence when bound to DNA and RNA. To help improve the specificity of ThT to amyloid, Kelenyi modified the staining conditions to a lower pH (between 0.8-2.8) with success (Kelényi 1967).

It is important to note that the inter-strand spacing along the fibril axis of 4.8 Å is unchanged in insulin fibrils upon ThT binding. This indicates that ThT binding is exclusive to the surface of amyloid fibres as an increased β-strand distance would have been evident in X-ray fibre diffraction data if ThT had bound between the β-strands (Groenning et al. 2007). Krebs *et al*. have since used confocal microscopy to show that the dye dipole excitation axis lies

parallel to the long molecular axis of the fibre. In other words, ThT binds to amyloid fibres such that their long axes are parallel *i.e.* in the 'channels' that run along the length of the cross β-sheets (Krebs, Bromley, and Donald 2005).

It has been proposed that ThT may bind to fibres in the form of micelles. A micelle is a form of molecular aggregation where various identical molecules, containing both a polar and a hydrophobic end, form a shell. The hydrophobic ends of each molecule are enclosed within the shell with the polar ends on the edges. In ThT the hydrophobic end consists of a dimethylamino group attached to a phenyl group and the polar end consists of a benzothiazole group. In 2005 Khuana *et al.* discovered that the fluorescence excitation and emission of ThT were dependent on the micelle formation. They noticed a correlation between micelle formation and transmission. The more micelles present the more transmission observed. The micelles were visualised using atomic force microscopy. They were 3nm in diameter and were bound proportionally along the length of the fibres. Upon disruption of the micelles using a pH of 3 or lower, binding of thioflavin T to amyloid fibres was also reduced. This suggests that the positive charge on the thioflavin T molecule has a role in its micelle formation that then bind the amyloid fibres (Groenning 2009). Although an interesting theory, the micelle proposal is inconsistent with the fact that ThT prefers to bind parallel to fibres. Perhaps the micelle theory is limited by ThT concentration or alternatively can be interpreted as stacking of ThT rotated along its lateral axis.

It is important to note that when looking at the level of fluorescence or fluorescence intensity on a spectrophotometer, the values obtained from one machine cannot be directly compared to those from another. This is because the resulting values are determined by the software, the detector systems, the electronics and filtering, etc of the machine in question. Due to this, ThT fluorescence in relation to cross β-sheet fibres is not given units. With that said, although the arbitrary values are not meaningful, they can be used for relative, ratiometric quantification of results obtained from one particular machine.

Congo Red

Congo red is a palindromic structure which demonstrates a characteristic apple green birefringence under polarised light when bound to amyloid. When examined between crossed polarizer and analyzer, a characteristic shift in absorbance maximum from about 490 nm to about 540 nm is seen. Congo red is not well suited for in situ detection, since Congo red interferes with processes of protein misfolding and aggregation, and is reported either to inhibit or enhance amyloid fibre formation for several proteins. Instead, Congo red is most often applied in diagnosis of amyloid in ex vivo tissue sections. This is usually done using polarization microscopy, but light microscopy and fluorescence microscopy can also be used. (Groenning 2009) **(Figure 14, Figure 15)**.

The use of Congo red as an detector of amyloid was first established in the 1920s by Benhold and Divry (Benhold 1922) (Divry 1927). They originally thought that the dye bound fibres in a crystal-like fashion with colour distribution being an indication of thickness. More recent work by Khurana *et al.* has shown that Congo red binds parallel to the β strands along the edges or the middle of the sheet (Khurana et al. 2001). They achieved this result through the use of circular dichroism (CD). Congo red binding was carried out with various proteins of different secondary structure classes. Throughout the course of their experiment they noticed that their CD bands were different from the tissue-extracted amyloid described by Benditt *et al.* (Benditt, Eriksen, and Berglund 1970). They concluded that these differences were most likely due to “contaminants” such as glycosaminoglycans and serum amyloid protein present in the *ex vivo* extracts.

Khurana *et al.* also noted that Congo red was able to bind to proteins in their native state. Pectate lyase, carbonic anhydrase (CA), β-lactoglobulin, IL-2, and lysozyme all demonstrated this however, after binding to the dye, each protein proceeded to oligomerise leading to the idea that stable Congo red binding is exclusive to multi-meric structures. In addition, CD peaks were 10

times smaller in the native bands than their fibre counterparts indicating that native proteins have fewer binding sites for the dye.

The binding of Congo red to such a variety of proteins in various forms indicates that amyloid binding is not structure specific for Congo red. Perhaps Congo red binding is limited to specific hydrophobic and electrostatic interactions which are more accessible in the fibre structure compared to the native counterparts. Most studies suggest that ionic interactions are important for Congo red binding to amyloid fibres. For example, the stoichiometry of Congo red binding decreases as pH increases in the binding of β-poly-L-lysine (Klunk, Pettegrew, and Abraham 1989). This suggests that protonation of the poly-L-lysine is important for binding. Furthermore, the importance of the ionizable group for binding of Congo red is demonstrated upon its removal. By removing one of these groups the affinity of Congo red to bind fibres is reduced dramatically (Cai, Innis, and Pike 2007).

Studies have also shown that hydrophobic interactions play an important role for burial of the hydrophobic part of the Congo red molecule (Wu et al. 2007). This would suggest that Congo red molecules are stabilized upon insertion into the grooves on the β-sheet surface. In addition, an increase in hydrophobicity shifts the absorption band of the dye to around 498 nm. This is similar to the shift observed in amyloid fibres but on a much smaller scale (Miura et al. 2002).

When used to analyse tissue samples, Congo red binding is not exclusive to protein let alone amyloid. Frequent binding to elastin, collagen fibres, RNA polymerase and cytoskeletal protein often results in false positive results (Woody, Reisbig, and Woody 1981) (Kagan, Hewitt, and Franzblau 1973). In an attempt to prevent this, tissue assays are performed under extreme conditions of 50–80% ethanol, high salt and alkaline pH conditions to yield binding to amyloid.

Although the mechanisms behind ThT and Congo red binding have been speculated various times in literature they are still not fully understood. The

one notion that *can* be concluded is that they bind to different sites on the amyloid fibre. This is derived from the fact that each has a different charge; ThT is positively charged whereas Congo red is negatively charged. Also, the binding kinetics of Congo red to amyloid fibres is slower than that of ThT suggesting that Congo red may not have easy access to its binding site most likely due to its size (Nilsson 2004).

It should also be noted that commercial stocks of these dyes are not always 100% pure (in accordance with the Sigma Aldrich website). ThT has a reported purity of only 65–75% with its main impurities being salts containing sodium, chloride, and sulphur. Commercial Congo red is contaminated with sodium chloride and water, making quantification by weight inaccurate. Recrystallization is likely to improve the purity of each and make quantification by weight possible. However, many authors seem to be unaware of this so it is difficult to directly compare their results.

Both ThT and Congo red are regularly used as a diagnostic tool but neither is 100% accurate or practical. Not only are both prone to bind bacteria and other tissue matter, but they can cause patients to go into anaphylactic shock if injected intravenously. This is a severe form of allergic reaction with symptoms such as swelling facial features, rash, wheezing, unconsciousness, nausea and vomiting. Due to this, other methods have been developed in order to analyse amyloid fibres. These include transmission electron microscopy (TEM), circular dichroism (CD) and x-ray crystallography.

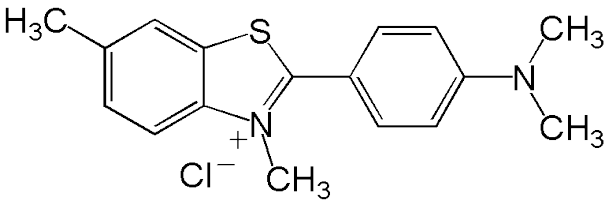


Figure 13: Diagram of Thioflavin T (ThT). It is constructed of a hydrophobic end and a polar end. The hydrophobic end consists of a dimethylamino group attached to a phenyl group and the polar end consists of a benzothiazole group. ThT is thought to bind proportionally to the channels which run parallel to the β-sheet. When bound ThT produces fluorescence at wavelength 420nm.

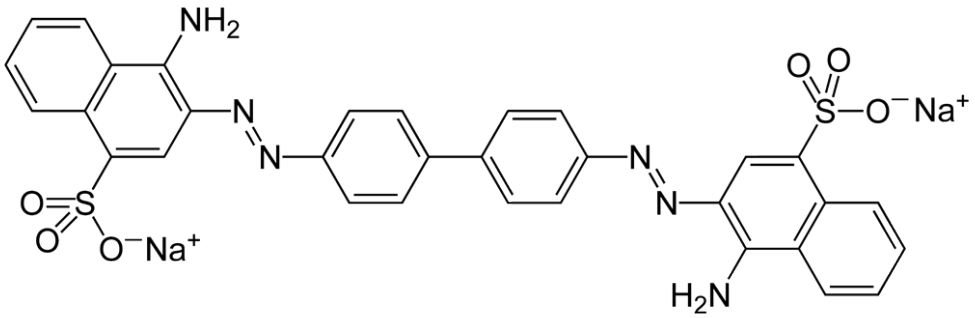


Figure 14: A diagram of Congo red. Congo red is a palindromic structure which binds parallel to the β strands along the edges or the middle of the sheet using both hydrophobic and electrostatic interactions. A characteristic apple green birefringence is demonstrated under polarised light when Congo red binds to amyloid.

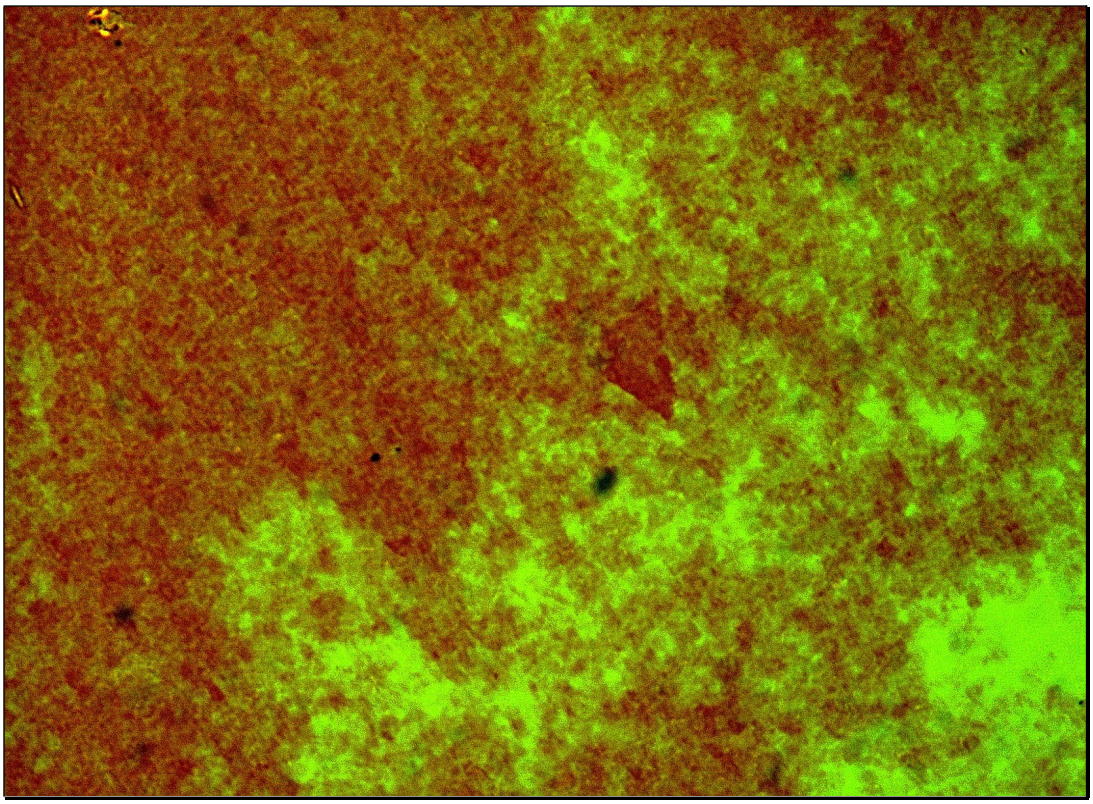


Figure 15: A photograph of insulin fibres grown in TE buffer and stained with Congo red. The sample was viewed under polarised light in order to see the apple green birefringence typical of cross-β sheets. Image was viewed using a Zenith Microlab ML-1000M Monocular Laboratory Microscope 60040 at a magnification of x40.

Transmission Electron Microscopy

Transmission Electron microscopy (TEM) is the gold standard for visualising amyloid fibres as it provides an actual photograph of the structures. TEM is a microscopy technique that uses electrons as a way to illuminate thin films of samples similar to the way optical microscopes use light (Figure 16).

In order to carry out TEM, the specimen or sample must exist in one of two forms: either as an ultrathin section less than 100 nm thick or as a suspension on a copper grid. The reason behind such thin samples is to allow the beam of electrons to transmit through the sample and interact with it. Electromagnetic fields focus the electrons in a way analogous to a lens in a light microscope however, unlike light in regular microscopes, the beam of electrons must be contained within a vacuum. This is due to the fact that electrons can easily be absorbed by air at atmospheric pressure. The use of electrons allows higher magnifications to be achieved in comparison to light microscopes. Electrons have smaller wavelengths than light so are able to interact with smaller objects. The image is then magnified and focused onto an imaging device, such as a fluorescent screen, a layer of photographic film, or a sensor such as a charge-coupled device.

Although a reliable way to identify amyloid fibres, TEM is a specialised form of analysis. The microscope itself is expensive and its use requires extensive training (D. B. Williams and Carter 1996) (J. C. Williams and Paton 1978).

Circular Dichroism

Circular Dichroism (CD) is an alternate technique which can be used. It works by using polarised light as a way to analyse chiral molecules of all types and sizes. It is a favoured technique as it is non-denaturing and has the ability to take into consideration any sensitivity of the sample towards environment, temperature and pH.

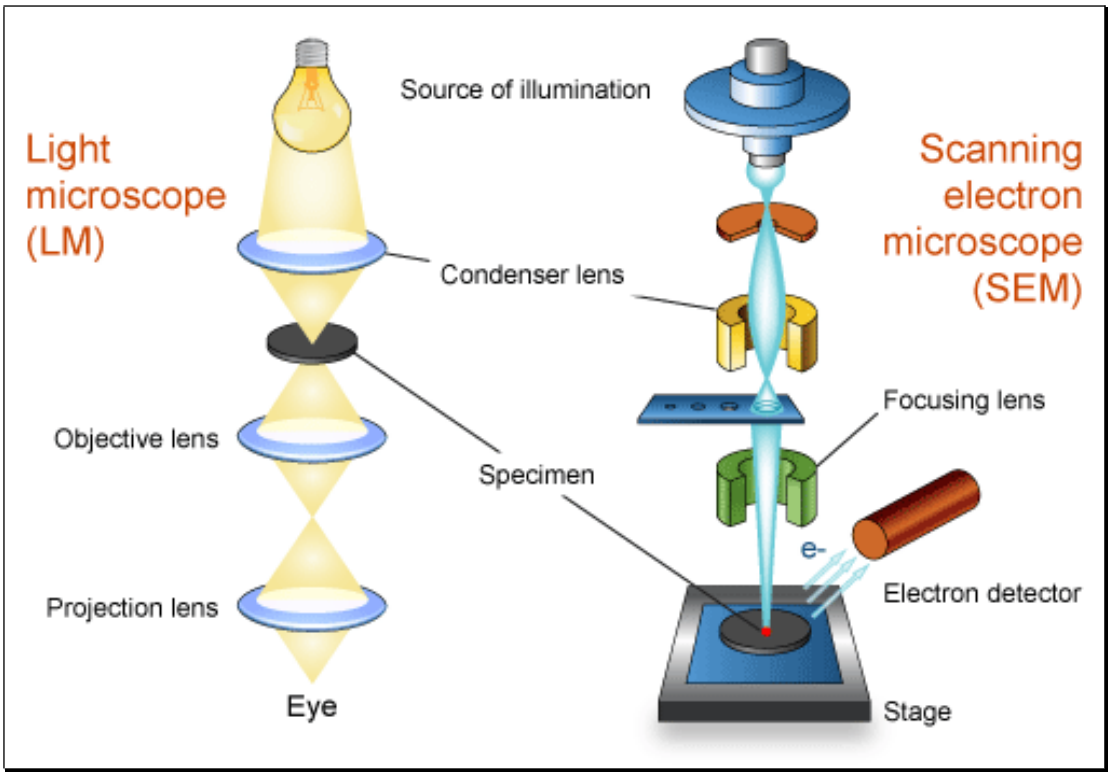


Figure 16: Cartoon comparing a light microscope with a transmission electron microscope (TEM). In a light microscope, the source of illumination is optical light. The light waves are condensed, focused and projected using glass lenses and the final image is viewed by the naked eye. It has a magnification of 1000-2000x which is a physical limit imposed by the wavelength of light. In comparison, in TEM the source of illumination is electrons. They are contained within a vacuum and are focused via magnets. An image is created on a computer screen when the electrons reflect from a metal coated sample and are detected by an electron detector. TEMs have a much higher resolution than light microscopes – up to 2 million times – however they are very expensive and require special training to use.

The need for chirality in CD makes it an ideal way to analyse proteins as they themselves are large chiral molecules (*i.e.* non super-imposable molecules). Chiral molecules exist as pairs of mirror-image isomers which are known as enantiomers. The physical and chemical properties of a pair of enantiomers are identical with two exceptions. They interact with polarised light differently and they also interact with other chiral molecules differently. It is the first of these exceptions which is taken advantage of in CD techniques.

In order to carry out CD, one must have two circularly polarised lights (CPL) of equal magnitude; one rotating clockwise and the other rotating anti-clockwise. These opposing circularised lights are passed through a solution of the optically active medium and measurements are carried out in the visible and UV region of the electromagnetic spectrum over a range of wavelengths. A spectopolarimeter monitors the electronic transitions. If the molecule under study is indeed chiral, then one CPL state will be absorbed to a greater extent than the other. This causes the CD signal over the corresponding wavelengths to be non-zero (Atkins and Paula 2005) **(Figure 22)**.

It is this difference in absorbance between the two wavelengths which is used as an indicator of the secondary structure of the protein (Kelly, Jess, and Price 2005). A CD signal can be positive or negative, depending on whether the anti-clockwise CPL is absorbed to a greater extent than the clockwise CPL (CD signal positive) or to a lesser extent (CD signal negative).

In order to truly understand circular dichroism, one must first understand the basics of polarisation beginning with unpolarised light. A light wave is a transverse electromagnetic wave. It is produced by oscillating electric charges and has both an electric and a magnetic component. When travelling through space, although travelling in a linear direction, the electric and magnetic oscillations of an electromagnetic wave occur in numerous planes. A light wave that is vibrating in more than one plane is referred to as unpolarised light**.** Linearly polarised light occurs when these oscillations are confined to a single plane. This can be achieved through a number of ways including

transmission, reflection, refraction and scattering. In CD however, polarisation occurs through use of a polaroid filters *a.k.a.* wave plates (Singh 2010).

Wave plates are made of a special material that is capable of blocking all but one of the planes of vibration of an electromagnetic wave. When unpolarised light is transmitted through a wave plate, it emerges on a single plane. A wave plate is able to polarize light because of the chemical composition of the filter material. The filter can be thought of as having long-chain molecules that are aligned within the filter in the same direction. As unpolarised light strikes the filter, the portion of the waves vibrating parallel to the long chain molecules are absorbed by the filter. The alignment of these molecules gives the filter a polarization axis. This polarization axis extends across the length of the filter and only allows vibrations of the electromagnetic wave that are parallel to the axis to pass through. Any vibrations that are perpendicular to the polarization axis are blocked by the filter.

To simplify matters, light waves are visualised as having their coupled oscillating electric field and magnetic field as perpendicular to one another. All polarised light states can be described as a sum of those two linearly states at right angles to each other, usually referenced as vertically and horizontally polarised light. If for instance we take horizontally and vertically polarised light waves of equal amplitude that are in phase with each other, the resultant light wave is linearly polarised at 45 degrees. Circularly polarised light occurs when the two polarisation states are out of phase. For example, if one of the polarised states is out of phase with the other by a quarter-wave, *i.e.* one of the linear components of the beam is slowed down with respect to the other so that they are one quarter-wave out of phase, the resultant will be a helix and is known as circularly polarised light (CPL). The helices can be either right-handed (R-CPL) or left-handed (L-CPL) and are non-superimposable mirror images. On traversing the solution of chiral medium, the phase relationship between the circularly polarised waves changes and the resultant linearly polarised wave rotates. This is the origin of the phenomenon known as optical rotation, which is measured using a polarimeter (Avadhanulu and Kshirsagar 2014).

There are both advantages and disadvantages for using circular dichroism to determine the presence of amyloid. It is a rapid, non-destructive technique which has the ability to determine native structural features, stereochemistry and can measure exchanging structures in changing environments. However, it cannot give structural information at the atomic level, it requires relatively concentrated (0.5mg/ml) sample solutions in order to produce results and it cannot be used for mixed sample solutions. It is limited to a single protein at a time.

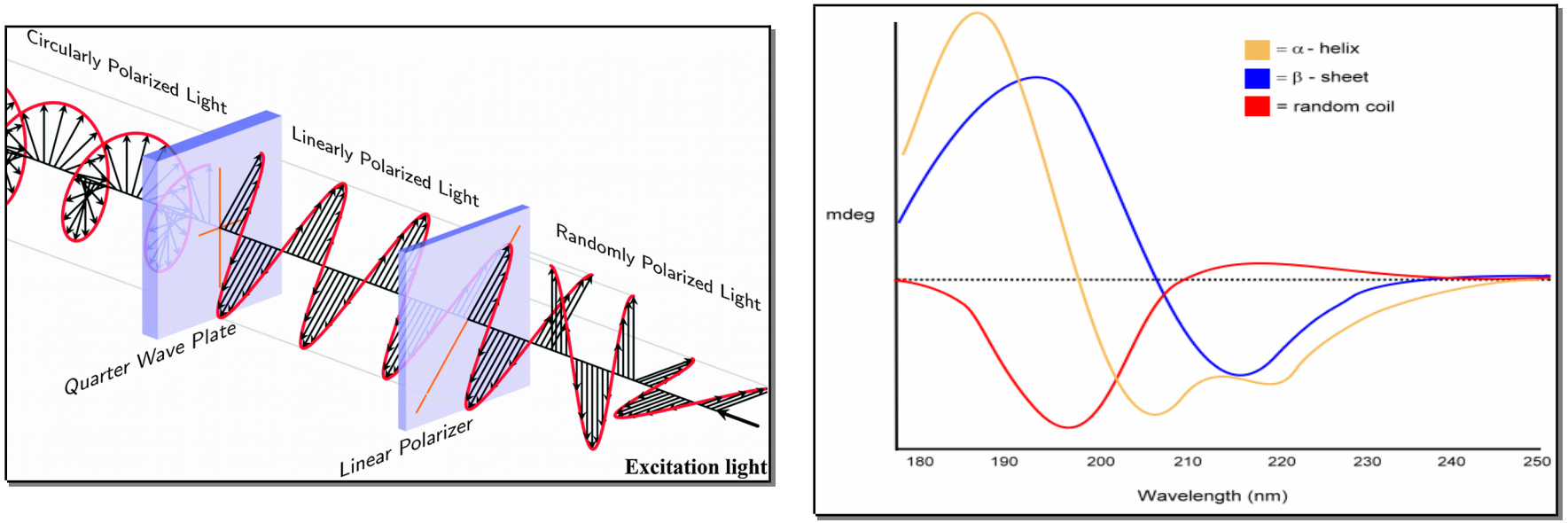


Figure 17: (Above) A graph showing the typical Circular Dichroism (CD) spectra for alpha helices (gold), beta sheets or strands (blue) and random coiling (red). The alpha helices produce a spectra consisting of a peak and double depression in the positive and negative mdeg axis respectively. The beta sheets/strands produce a spectra consisting of a peak and single depression in the positive and negative mdeg axis respectively with a shifted x axis interception point to the right. Random coils produce spectra consisting of a depression followed by a low but broad peak. (Below) Image depicting the creation of circularly polarised light. The electric and magnetic oscillations of an electromagnetic wave occur in numerous planes in un-polarised light. Linearly polarised light is obtained in CD when these oscillations are confined to a single plane through use of a wave plate. When two polarisation states are out of phase with one another circularly polarised light is formed (Chu et al. 2015).

X-Ray Crystallography

X-ray crystallography is the most difficult method used for analysing amyloid. In this technique the atoms of the sample are placed in the path of an x-ray and this obstruction causes a beam of incident X-rays to diffract into many specific directions. The refracted beams are monitored using a detector and processed into a pattern of spots. The dispersion and intensity of the spots are measured by a detector are later used by a computer to determine the atomic and molecular structure of the sample. – Further details in Chapter 4: "X-Ray Crystallography"

The reason this method is described as difficult is because it requires the sample to be in crystalline form. Turning a sample into a crystal is an extremely time consuming process. It requires various steps, each of which must occur successfully in order for a crystal to form. Firstly, the desired protein must be successfully expressed and purified through use of prokaryotic or eukaryotic recombinant techniques. Once achieved, protein crystals must be grown to an adequate quality and size, harvested and frozen without causing any damage to the fragile structure. Finally, diffraction must occur at a high enough resolution in order to reveal atomic detail. Each of these steps is incredibly time consuming, may not necessarily work, and can require great amounts of trial and error (G Wagner, S G Hyberts, and Havel 1992).

TEM, CD and x-ray crystallography methods are all frequently used for *in vitro* studies of amyloid fibres and are often referred to in the literature. They are each a more reliable method of detecting and analysing amyloid compared to Congo red and ThT dyes. However, they are also more expensive, require more skill to administer and take far longer to complete. The use of Congo red and ThT dyes is quicker and easier which explains their continued use today as a research and diagnostic tool.

Scintigraphy

There is one final method used for analysing amyloid called scintigraphy. Scintigraphy is slightly different to the previous examples as it is performed *in vivo* **(Figure 18)**.

Scintigraphy involves using a scintillation counter alongside a radioactive tracer in order to obtain an image of a bodily organ or a record of its functioning. In the case of amyloidosis, the protein Serum Amyloid P (SAP) is used as the radioactive tracer. SAP is a 125kDa pentameric protein with an affinity for amyloid. It is found in the serum and has a half life of 1 day. In 1988 Hawkins *et al* performed *in vivo* studies involving iodine-123 labelled SAP. These involved injection of the radio-labelled SAP into affected patients and monitoring its distribution through use of a gamma camera. Those suffering from amyloidosis demonstrated an accumulation of the radio-labelled SAP preventing a clearance of it from the body and increasing its half-life from 1 to 30 days (Hawkins, Lavender, and Pepys 1990; Hawkins et al. 1988). Sørensen *et al*. in 2000 proved that these interactions were calcium dependent by demonstrating that complexes could not be formed in plasma containing EDTA (a metal chelator).

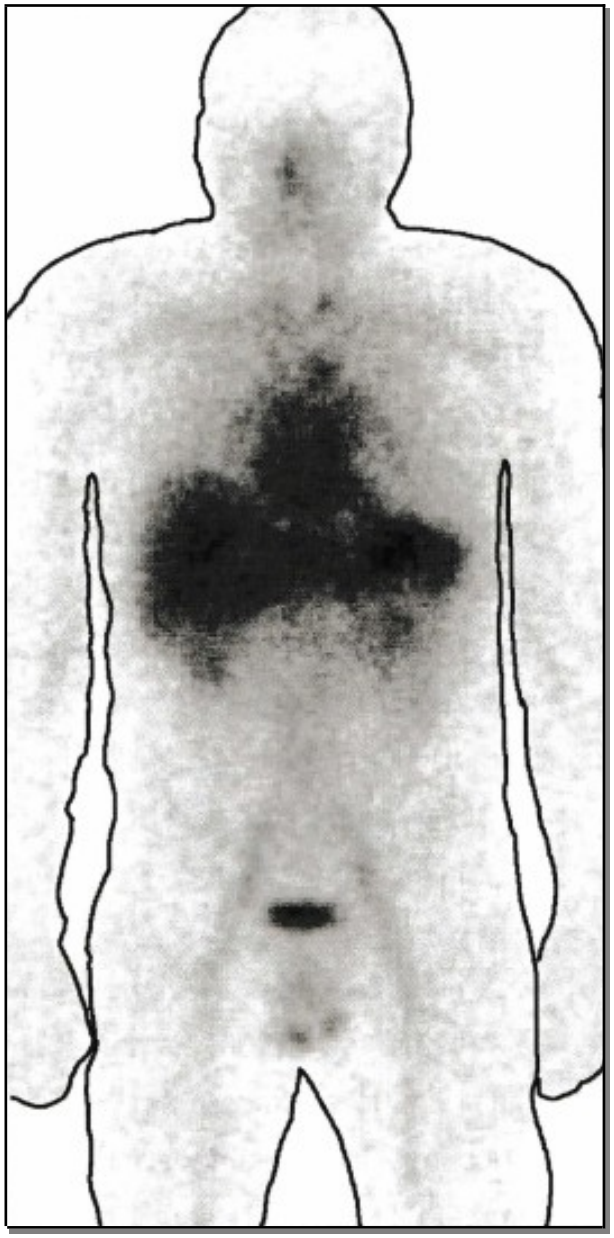


Figure 18: Radio-labelled SAP scans from a patient suffering from AL amyloidosis. Iodine-123 labelled SAP was injected and monitored via a gamma camera. Darkened areas represent SAP accumulation - i.e. amyloid deposit.

An Introduction to Serum Amyloid P Component (SAP)

Serum amyloid P component (SAP) is referred to as a universal component of Amyloid. Various macromolecules have the ability to bind to amyloid fibres to form the structures referred to as amyloid deposits, however, the majority of these macromolecules bind specifically to amyloid fibres of a particular protein. SAP and glycosaminoglycans are the only components universal to all deposits regardless of the protein involved (Ancsin 2003) (Skinner and Cohen 1988). The roles of SAP and glycosaminoglycans are unknown, however it is believed that they are involved in protecting amyloid against degradation and may even enhance fibril folding (Tennent, Lovat, and Pepys 1995).

SAP is a plasma glycoprotein of unknown function which makes up to 15% total mass in all amyloid deposits (Calkins 1983). It is made in the liver and travels in the plasma at a concentration of 30-45mg/L (Rubio et al. 1993). SAP belongs to the pentraxin family of proteins and like the others in this family, exhibits calcium dependent ligand binding. It has a ring/disc-like structure consisting of 5 identical subunits arranged non-covalently in pentameric radial symmetry and is often referred to as an SAP pentamer (M. B Pepys et al. 1977) **(Figure 19)**.

SAP Structure

SAP crystals were first observed in 1986 however crystallisation proved to be difficult due to SAP's tendency to precipitate in the presence of calcium ions. Fortunately this was overcome by Wood *et al* in 1988 through the use of calcium site ligands and the structure was finally published in 1994 by Jonas Emsley (Jonas Emsley 1994). The resulting crystal structure shows a symmetrical disc (pentamer) 100Å in diameter, 35 Å in depth with a 20Å hole in the middle.

Figure 19: Ribbon diagram of a SAP pentamer (A-Face). Loops are represented in white, β-strands in gold and α-helices in magenta. The resulting crystal structure shows a symmetrical disc (pentamer) 100Å in diameter, 35 Å in depth with a 20Å hole in the middle. The 5 subunits interact non-covalently through the N and C termini of adjacent subunits via Van der Waals and hydrogen bonding. Each identical SAP subunit is 204 residues long, 25kDa in size and folded into 2 layers of anti-parallel strands which sandwich a hydrophobic core between them. The tertiary structure is stabilized through disulphide bridge linking of cysteine residues 36 and 95 (Ozawa et al. 2016) (PDB code: 1SAC).

Each identical SAP subunit is 204 residues long, 25kDa in size and contains a sequence signature core common to all pentraxins (HxCxSx/TWxS) (Handbook of animal lectins: properties and biomedical applications 2000). The pentraxin domains are of jellyroll topology and consist of the following components: ß-sheets, ß-turns, unknown structures (*e.g.* random coiling) and α-helixes to the ratio 54:25:19:2 respectively (Dong, Caughey, and Du Clos 1994). These components fold into 2 layers of anti-parallel strands which sandwich a hydrophobic core between them containing the residues leucine 62, tyrosine 64 and tyrosine 74 (Srinivasan et al. 1994). This tertiary structure is further stabilized by the disulphide bridge linking cysteine residues 36 and 95, as well as the N- linked attachment of a single biantennary oligosaccharide attached to asparagine 32 (Hohenester et al. 1997) (Alun W Ashton 1997)**.** The 5 subunits interact non-covalently through the N- and C-termini of adjacent subunits via Van der Waals and hydrogen bonding (Dong, Caughey, and Du Clos 1994) (Mark B. Pepys 2006) **(Figure 20,)**.

SAP Calcium Binding

One side of each subunit is flat and contains an α-helix and is so named the A-face whereas the other side, known as the B-face, is twisted and has a crevice created by neighbouring loops. It is this B-face crevice which is responsible for calcium ion binding (Kinoshita et al. 1992).

A total of 2 calcium ions are able to bind to each subunit leading to a maximum of 10 calcium ions per SAP pentamer. The calcium ion pairs bind 4.2Å apart and are stabilised via interactions between carboxylate and amide residues in the binding site (Jonas Emsley 1994). The first of each pair of calcium ions binds via interactions with aspartate 58, asparagine 59, glutamic acid 136, glycine 137 and aspartic acid 138, and the second with glutamic acid 36, aspartic acid 138 and glutamine 48 in a water mediated hydrogen bond (Gewurz, Zhang, and Lint 1995). In the absence of calcium, some enzymes such as pronase and α-chymotrypsin are able to cleave the exposed calcium loop causing the formation of a less stable structure (István Likó

2007). Experiments have shown that the binding of calcium ions leads to the SAP structure becoming very stable and highly resistant to proteolysis (Tennent, Lovat, and Pepys 1995). This is thought to explain why amyloid fibres are so resistant to degradation when bound by SAP **(Figure 22)**.

SAP Conformations

The SAP ring has polarity due to the presence of the calcium binding sites on its B-face (Hohenester et al. 1997). Crystal analysis has shown that in the absence of calcium and at a pH of 5.5 or higher, two individual SAP molecules have the ability to interact through ionic interactions to form a B-face to B-face decamer (O'Hara et al. 1988). Thompson *et al.* scrutinised this further and demonstrated that SAP actually exists as a mixture of pentamers and decamers in these conditions; SAP does not exclusively form decamers (D Thompson 2002).

That being said, one would expect SAP to exist as a decamer mixture in the serum. Serum has a physiological pH of 7.35-7.45 with a low calcium concentration of 8.5-10.2mg/dL of which only 4.8-5.7mg/dL exists as free ionised ion in serum. This is not sufficient to saturate the SAP when their diverse molecular masses are taken into consideration. However, Sorensen *et al.* disproved this in 1995. They purified fresh samples of human SAP and analysed them via various methods including permeation chromatography, quantitative immunoelectrophoresis, Western blotting, and electron microscopy. Their results showed that SAP travels in the blood exclusively in the pentameric form and not in the decameric form expected (Sørensen et al. 1995). The reason behind this is uncertain however their data indicated that native SAP may form a complex with other serum proteins such as C4b-binding protein. In doing so it is unable to interact with another of its kind.

The fact that the decamer is only observed *in vitro* shows that it is probably caused from ligand mediated contact or from electrostatic interactions between the 2 faces (Hutchinson, Hohenester, and Pepys 2000). It was

originally thought to be formed by A-face to A-face interactions. Scientists believed this because SAP decamers are susceptible to cleavage at the calcium loop located on the B-face. This loop was only thought to be exposed in an A-face interaction. Recent crystallographic data, however, has shown that the loop is actually on the outer edge of the B-face and, subsequently, does not obscure B-face interactions (Alun W Ashton 1997).

SAPs tendency to auto-aggregate is a common problem during *in vitro* experimentation. Crystal forms of SAP suggest that in the case of auto-aggregation, an A-face to B-face interaction occurs instead. Here, glutamic acid 87 on the α-helix N-terminus of one pentamer protrudes towards the calcium binding site of the B-face of the adjacent pentamer. This issue can been resolved through the use of calcium dependent ligands and supra physiological concentrations of salt. Additionally, point mutagenesis of this residue to either a glutamine or a serine abolishes auto-aggregation. In the cases where calcium dependent ligands are involved, glutamic acid 87 is unable to interact with a neighbouring pentamer as the ligand is occupying the calcium binding site. Kolstoe's studies in 2005 showed that in high salt and calcium concentrations it is the chloride aspect which is most likely preventing the glutamic acid 87 interaction. The reason in which the concentrations need to be so high is so to ensure full saturation of all five calcium binding sites present on each pentamer (Kolstoe, S.E. 2005)

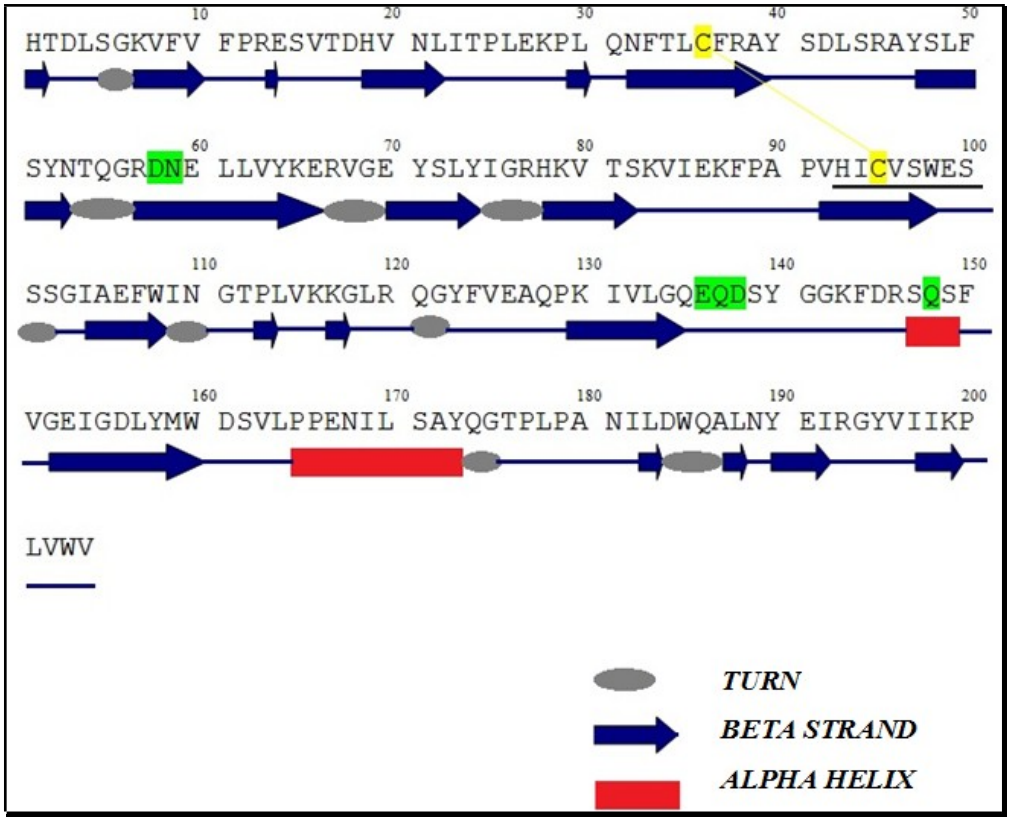


Figure 20: The primary sequence of an SAP domain with its associated secondary structures marked: ß-strands, ß-turns, and α-helixes to the ratio 54:25:2 respectively. These components fold into 2 layers of anti-parallel strands which sandwich a hydrophobic core between. This is stabilized by a disulphide bridge linking cystein residues 36 and 95 (highlighted in yellow), as well as an N- linked attachment of a single biantennary oligosaccharide to asparagine 32. The residues involved in calcium binding are highlighted in green and the pentraxin signature sequence common to all pentraxins is underlined in black.

Figure 21: Ribbon diagram of an SAP domain A-face (top) and B-face (bottom). The β-strands are labelled A to O from the N- to C- terminus and are folded into a jellyroll topology. The α-helix present on the A-face is highlighted in green and the stabilising cysteines 36 and 95 are represented in yellow (Jenvey, 2006).

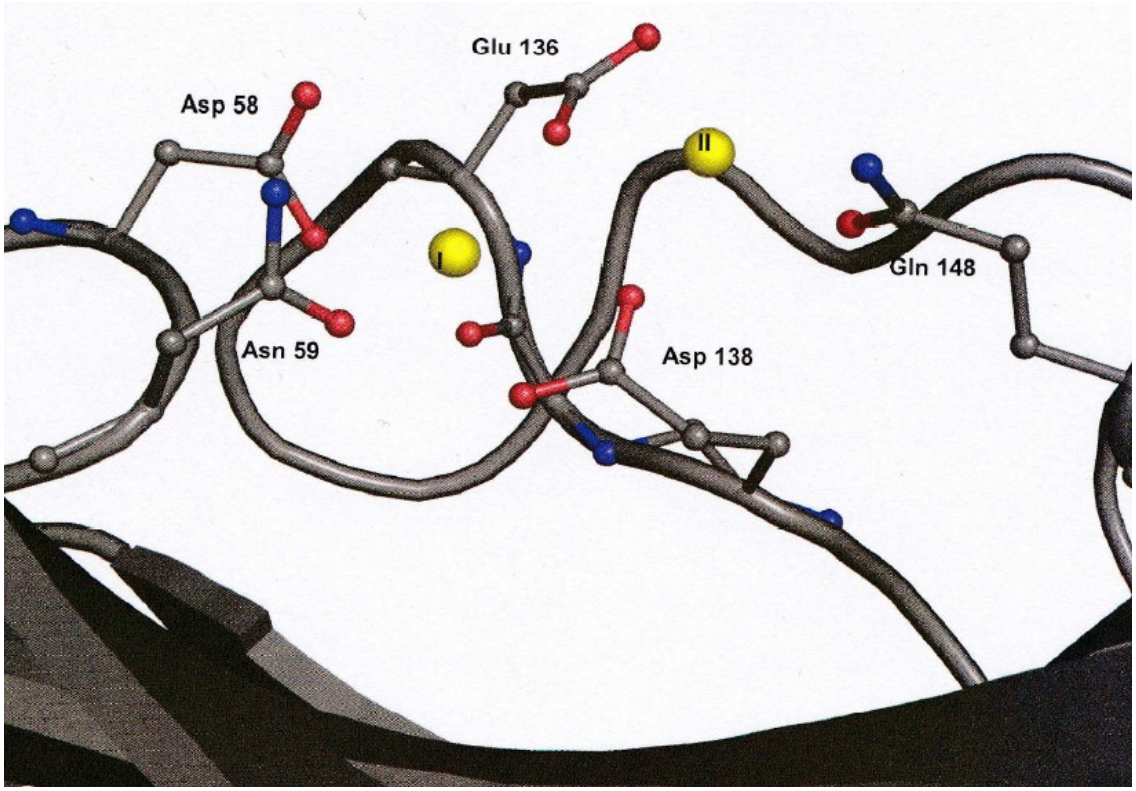


Figure 22: Calcium binding site of SAP. Calcium ions are represented in yellow. A total of 2 calcium ions are able to bind to each subunit leading to a maximum of 10 calcium ions per SAP pentamer. The calcium ion pairs bind 4.2Å apart and are stabilised via interactions between carboxylate and amide residues in the binding site (Jonas Emsley 1994). The first of each pair of calcium ions binds via interactions with aspartate 58, asparagine 59, glutamic acid 136, glycine 137 and aspartic acid 138, and the second with glutamic acid 36, aspartic acid 138 and glutamine 48 in a water mediated hydrogen bond (Gewurz, Zhang, and Lint 1995).

SAP Function

Despite the structure of SAP being well defined, its physiological role is still unclear. SAP has many ligands including DNA, chromatin, C1q and PE but their relevance has not been determined.

SAP is the only serum protein which specifically binds to DNA and long, native chromatin and is able to displace the H1 histones (M. B. Pepys and Butler 1987). Histones are responsible for compressing and packaging of DNA into the cell nucleus. In vitro studies have shown that their displacement results in the solubilisation of chromatin in buffers of physiological strength. In 1992 Hicks et al discovered that SAP actually has the ability to bind directly to histones with the resulting complex activating the classical pathway of compliment. In doing so it allows antibodies and phagocytic cells to remove phagocytes from the body (Hicks et al. 1992). SAP has also been known to bind chromatin in vivo. This is seen, in particular, in biopsies from patients with systemic lupus erythematosus (SLE); an autoimmune disease in which the body's immune system mistakenly attacks healthy tissue. In this case, the plasma concentration of SAP remains at 30mg/L regardless of an immune response indicating SAP does not have an immunological role (Breathnach et al. 1989).

SAPs affinity for both single and double stranded DNA suggests that it could be involved in the removal of apoptotic debris and possibly have an anti-inflammatory role. When cells die *in vivo* the leakage of plasma proteins triggers an immune response in the form of leukocyte recruitment. An immune response is necessary in such situations as it alerts the host to "danger zones" and prevents the pathogenesis of a number of diseases. The increased vascular permeability of the "danger zone" very rapidly increases the delivery of blood borne defences such as antibodies to the affected site in an attempt to clear them. SAP, being an extracellular protein, would never have access to DNA other than after cell death so an anti-inflammatory role is plausible considering its affinity for DNA (Rock and Kono 2008) (Sipe 2006).

Another explanation for this interaction could be to prevent anti-DNA antibodies forming such as those seen in SLE. This was shown in a knockout mouse study by Bickerstaff et al 1999. The mice accumulated pathogenic anti-nuclear antibodies to DNA, chromatin and even histones. Also, a high proportion of test subjects developed proliferative glomerulonephritis - a phenotype reminiscent of human SLE (Bickerstaff et al. 1999).

The Involvement of SAP in Amyloidosis

In 1987 Hamazaki demonstrated that SAP binds specifically to the heparin and dermatan sulphate components present on the surface of amyloid deposits. These interactions occur in a calcium dependent manner under physiological conditions (Hamazaki 1987). In 1995 however, Tennent et al showed that SAP is able to bind amyloid formed from pure proteins and peptides *in vitro* with the same affinity. This indicates that SAP has a binding affinity to the structural motif of amyloid fibres, not to the substrates bound to them. Upon attempting to bind SAP with those proteins in their native form (insulin, TTR and Aβ), Tennent found that binding did not occur. This demonstrates that this structural motif is exclusive to amyloid – *i.e.* the cross β-sheet (Sipe 2006).

As previously discussed, SAPs role and function is yet to be fully established. Nonetheless, certain behaviours have been observed *in vivo* and *in vitro* in relation to amyloid fibres. *In vivo* studies have shown SAP to have an anti-opsonin role. The immune system is normally able to recognise amyloid but SAP provides a protective coat around the fibres. This in turn prevents phagocytic cell recognition and thus prevents an immune response in the presence of amyloid deposits (O'Nuallain and Wetzel 2002). There have also been other studies demonstrating this 'pro-fibre' behaviour of SAP such as those carried out on SAP knock-out mice. Here, the genetically altered mice exhibit a retarded deposition of amyloid and reduced total fibre mass (Botto et al. 1997). This would imply that SAP is involved in catalysing or enhancing amyloid formation.

In vitro, Hamazaki, Webster and Rogers have been able to demonstrate SAPs fibrillogenic nature using amyloid-β (Aβ) fibres. They observed that fibre concentrations increased upon the addition of SAP. These observations, alongside the fact that SAP bound fibres are more stable and resistant to proteolysis *in vitro* suggests that SAP has a role as a fibrillogenic stabiliser or enhancer (Hamazaki 1987).

Contrastingly, Janciauskiene *et a*l. in 1995 observed the opposite effect; SAP actually prevented Aβ fibres from forming *in vitro*. They also noted that SAP has the ability to bind amyloid fibres made out of any protein but could not bind to these proteins in their native state suggesting that as well as its fibrillogenic involvement, SAP may also have a role as an amyloid molecular chaperone. These results began a set of debates between the two lab groups. Neither could understand why these opposing results were being produced.

Upon closer inspection one can see that there is a very important difference between the two experimental methods. Hamazakis group performed their assays in the presence of a physiological concentration of Ca^{2+} whereas Janciauskienes group performed theirs in the absence of any calcium ions. This introduction has already established SAPs ability to exist as either a pentamer or a decamer depending on its surrounding environment. In the absence of calcium and at a pH of 5.5 or higher, two individual SAP molecules have the ability to interact through ionic interactions to form a B-face to B-face decamer, or more precisely, as a mixture or decamers and pentamers. With this in mind, one can assume that SAP was in its pentameric form in Hamazakis assays and in its decameric form in Janciauskienes assays.

Experimental Aims

The overall aim of this thesis is to clarify the role of SAP in *in vitro* fibre formation. Does SAP behave as a fibrillogenic enhancer as suggested by Hamazaki or does it, in fact, decrease fibre mass as implied by Janciauskiene? The most likely answer is that SAP behaves in both ways and that each role is specific to a particular face or conformer. In order to investigate this, cross β-sheet fibres will need to be exposed to each SAP face/conformer independently and the behaviours of the fibres observed.

When in the pentameric state - a formation achieved in the presence of calcium - SAP will prompt interactions exclusively with its B-face. This is because in these conditions, the B side of SAP is bound to 10 calcium ions and exhibits calcium mediated binding as explained in the literature.

When in its decameric state - a formation achieved in the absence of calcium at a pH of 5.5 or higher - SAP should prompt interactions exclusively with its A-face. In these conditions the B-faces of two separate SAP molecules interact with one another and so are unavailable for other interactions. Thus far, there has been no indication in the literature as to whether A-face interactions are possible between SAP and fibres however this seems the most plausible explaination for the apparent dual behaviour of SAP.

Once the role of each face/conformer has been determined it would become necessary to find out whether this behaviour is applicable to other fibre forming proteins. This would establish whether SAPs behaviour is consistent among all amyloid proteins or just specific to one.

This thesis aims to conduct all its experiments at neutral pH. Normally, when dealing with amyloid fibres, highly acidic conditions are usually implemented; attempts have not been made at neutral pH. However, due to the inclusion of SAP in these experiments, it is important to use a neutral pH as the

denaturing conditions presented by a low pH would potentially alter SAP structure and function and therefore make results unreliable and incomparable. In addition, the stress introduced by the strong acidic conditions could also induce loss of SAP solubility and promote aggregation due to exposure of its hydrophobic groups. By using more favorable conditions the risk of altering SAP structure or function is removed and the results obtained are not compromised. Furthermore, at a neutral pH one would have a more physiological representation of SAPs effect on fibre formation.

In addition to the functional studies described above, this thesis also intends to carry out some structural work on the fibres and SAP molecule. TEM and CD analysis will be used to compare the morphologies of fibres made from different proteins and in different conditions. Although it is known that all amyloid fibres share the cross β-sheet motif regardless of the protein involved, it seems unlikely that they would all be completely identical in terms of morphology. There are more than 50,000 proteins in the human body alone, each of which has a unique amino acid sequence. It would be absurd to think that their individual sizes and chemical compositions would not have any influence over their cross β-sheet structure. Furthermore it is likely that SAP also has some effect on fibre morphology. It has already been established that the presence of SAP can increase fibre stability. Does SAP achieve this solely through binding to the fibres or does it have an influence on structure?

Finally, this thesis aims to look at the relationship between SAP and antibodies. GlaxoSmithKline have recently developed antibodies which are specific to SAP. These anti-SAP antibodies have the ability to detach SAP from amyloid deposits and thus leave the deposits exposed to the host's immune system. The exact mechanism behind this is not fully understood however clinical trials have shown that following the removal of deposit bound SAP from the body, deposits are dissociated and fibre mass is reduced. Although these results appear promising, more work needs to be carried out in order to guarantee safety and effectiveness of the treatment. As of yet,

these trials have shown a decrease in amyloid but they have yet to be successful in complete amyloid eradication (further discussions in chapter 5).

To this date, extensive structural analysis has not been carried out on the SAP:antibody complex. It is not known how or where they bind. By solving the structure, the knowledge on how to create or optimise amyloid treatment could then be deciphered.

Concluding Remarks

Interest in the topic of amyloid formation by proteins has increased dramatically in recent years. It has been transformed from a puzzling phenomenon associated with a small number of diseases into a major subject of study in disciplines ranging from material science to biology and medicine. Scientists are realising that amyloid is far more prominent in medicine than previously thought. It has been hidden from research due to its vast array of symptoms and its difficulty to differentiate from other diseases and as a result of this, the knowledge on how to treat it is not very well established.

Presently there is no general ‘cure’ or treatment for amyloidosis. The physical effects of amyloidosis can only be reduced through drastic actions such as surgery and transplantation. By understanding the factors that influence the conversion of proteins into the fibre state one can begin to create therapeutic strategies against it. For example, one could decrease the concentration of the aggregation prone species in the serum or stabilise the variant. It may even be possible to block the growth of fibres and/or enhance the housekeeping clearance mechanisms. Perhaps the supposed chaperoning activity of SAP could be used to eliminate amyloid fibres.

SAP is a critical component of all amyloid deposits *in vivo* and a key factor in preventing its removal from the body. By understanding the mechanism through which SAP interacts with fibres one could possibly devise a way in

which to control it or manipulate it. In doing so, it could be possible to reverse the symptoms of all 36 diseases with one simple solution.

Chapter 2

Insulin Fibre Analysis

Insulin Fibre Analysis

Chapter Aims

The aim of this particular chapter is to conclude whether SAP has two independent roles in amyloidosis with respect to a single protein; one pro-fibrillogenic and the other anti-fibrillogenic. If this is proven to be true, this chapter will then look to see if these roles are specific to a particular face or conformer of SAP. The resulting fibres will then be analysed using structural methods such as TEM and CD in order to see what affect, if any, SAP has had on their final morphologies.

In order to give a clear indication of the role of each conformer in fibre formation, cross β-sheet fibres must be exposed to each face/conformer independently. To enforce this, each experimental assay must be carried out in slightly different conditions. The reason for this being that each conformer requires different conditions in order for it to form. For example, to prompt interactions exclusively with the pentamer, calcium must be present in the solution. When in the pentameric state, the B side of SAP is bound to 10 calcium ions exhibits calcium mediated binding. In order to prompt interactions with the decamer, SAP must be in the absence of calcium at a pH of 5.5 or higher. When existing as a decamer, the B-faces of two separate SAP molecules interact with one another and so are unavailable for other interactions. Thus far, there has been no indication in the literature as to whether A-face interactions are possible between SAP and fibres however this seems the most plausible explanation for the apparent dual behaviour of SAP.

Taking these factors into account the following steps were proposed:

- Produce protein fibres *in vitro* at neutral pH
- Expose one set of fibres to SAP pentamers during growth.

- Expose another set of fibres to SAP decamers during growth.
- Leave one final set to grow without SAP interference to act as a control.
- Use dyes, electron microscopy and other analytical techniques in order to observe the effect of SAP (or lack of) on fibril formation and morphology.

By carrying out the above on a single amyloid forming protein one will get an indication of SAPs role to that specific protein. However, in order to create a general 'rule' and show SAPs role with relation to all proteins, analysis of its interactions will have to be carried out on multiple sets of fibres; each made from a different protein (further details in chapter 3).

It was vital that all fibre assays are performed at neutral pH rather than the acidic conditions normally implemented for fibre formation. This was due to the inclusion of SAP in these experiments. The denaturing conditions presented by a low pH would potentially alter SAP structure and function and therefore make results unreliable and incomparable. In addition, the stress introduced by the strong acidic conditions could also induce loss of SAP solubility and promote aggregation due to exposure of its hydrophobic groups. By changing the conditions from an acidic to a neutral environment the risk of altering SAP structure or function is removed and the results obtained are not compromised. Also, at a neutral pH one would have a more physiological representation of SAPs effect on fibre formation.

Due to the need to perform these experiments at neutral pH it was deemed important to choose a relatively cheap protein for experimentation. The lack of literature on neutral fibres would reduce the amount of guidance available for this work. As a consequence to this, creating an appropriate and successful protocol for would likely require trial and error. This can be both time consuming and costly as there is no guarantee things will work.

The first protein/fibres chosen to undergo SAP analysis was insulin. Insulin is a relatively inexpensive protein which demonstrates a high tendency to form amyloid fibres *in vitro*. Lab made insulin fibres have been reported many times in literature. One example of this is by Bouchard *et a*l. in 2000. This team used a combination of Fourier transform infrared spectroscopy (FTIR), CD, and electron microscopy (EM) to follow the transition of insulin from a helical structure to the rope-like twist associated with cross β-sheet structure. They achieved this by exposing the protein to deuterium chloride in a strong acidic environment of pH 2.3. The solution was heated to 70°C and incubated for up to 18 hours. In 2002 Jiménez *et a*l. produced the same results by incubating bovine insulin at 60° in pH 2, hydrochloric acid.

As of yet, insulin fibres have only been produced in acidic conditions. In order to create insulin fibres *in vitro* at neutral pH, an alternate protocol would have to be devised.

It was decided to use surface tension as a way to partially unfold the insulin from its native state into amyloid prone intermediates. Surface tension is the phenomenon which causes the surface of a liquid to act as an elastic sheet. This term is typically used when liquid is in contact with a gas such as the air although it can also be used to explain the relationship between water and oil. Surface tension is caused by various intermolecular forces such as Van der Waals. These intermolecular forces vary depending on the nature of the liquid (e.g. water vs. gasoline) or solutes in the liquid such as detergents.

In the case of water, the water molecules attract one another due to the water's polar property. The hydrogen ends of one molecule, which are slightly positive, are attracted to the slightly negative oxygen end of another. This causes the molecules to "stick" together. In a sample of water there are two types of molecules: those that are on the exterior of the bulk liquid and those in the interior. The interior molecules are pulled equally in every direction by neighbouring liquid molecules, resulting in a net force of zero. However, the molecules at the surface do not have molecules on all sides of them and therefore are pulled inwards. This makes it so that the energy state of the

molecules on the interior is much lower than those on the exterior creating internal pressure which forces liquid surfaces to maintain a minimal surface area. This in turn allows more molecules to have a lower energy state.

It takes a certain amount of energy to break these intermolecular bonds. Water in particular is one liquid known to have a very high surface tension value which is difficult to overcome. By creating a continuous movement of surface tension this should increase the internal pressure of the sample during rotation and create enough force to denature the insulin monomers and allow amyloid prone intermediates to form.

To ensure another variable is not accidentally introduced, it is important that all rotation assays are carried out in identical containers. By doing so the forces responsible for denaturing the protein are equal and samples are subject to identical conditions. The results of each assay will be monitored by ThT fluorescence. This method was chosen as the primary method due to its quick and easy use. However, due to its unreliability, other methods such as CD and TEM will be used in conjunction to ascertain that cross β-sheet fibres have indeed been formed.

An Introduction to Insulin

Insulin is a peptide hormone responsible for regulating normal blood glucose levels. It is synthesised in the β cells of the pancreatic islets of Langherhans and matures from the precursors pre-proinsulin and proinsulin. The way in which insulin maintains blood glucose levels is by facilitating cellular glucose uptake, regulating carbohydrate, lipid and protein metabolism and promoting cell division and growth through its mitogenic effects. It is important to note that insulin does not carry out these functions directly, but instead triggers various biological pathways depending on the glucose levels present at the time. For example, in high glucose levels, high levels of insulin are secreted which help to promote the formation of fatty acids and inhibits glucose production in the liver. Alternatively, in low glucose levels, pathways promoting the oxidisation of sugars are activated instead (Heilbrunn et al. 1958).

Insulin Function *in vivo*

Glucose is the principal stimulus for insulin secretion although other macronutrients, hormones, and neural input can modify this response. When glucose enters into the β-cell, *i.e.* when there is a high concentration of glucose, it is sensed by the enzyme glucokinase and phosphorylated into glucose-6-phosphate (G6P). This causes the generation of ATP and the closure of K^+-ATP-dependent channels. Closure of K^+-ATP-dependent channels leads to membrane depolarization and activation of voltage dependent calcium channels. The resulting increase in intracellular calcium concentration triggers insulin secretion (Soria et al. 2004).

In healthy individuals, glucose stimulated pancreatic secretion is biphasic and often described as being pulsatile. It responds to a stimulus such as glucose with an initial rapid phase of insulin secretion and is then followed by a less intense but more sustained release of the hormone (Bratanova-Tochkova et

al. 2002). The first phase of insulin secretion represents release of insulin already synthesised and stored in secretory granules. The second phase represents secretion of both stored and newly synthesised insulin (DeFronzo et al. 2015)

Insulin Synthesis

The gene for insulin is located on the short arm of chromosome 11 (Schröder and Zühlke 1982). It is synthesised in the ribosomes of the rough endoplasmic reticulum (RER) as a precursor called pre-proinsulin. Pre-proinsulin is a single chain, 100 amino acids in length, formed by sequential synthesis of the following components: a signal peptide, the B chain, the connecting (C) peptide and the A chain. The removal of the signal peptide by the signal peptidase SEC11 is the first step in insulin maturation. It shortens the chain to a structure 86 residues long referred to as proinsulin. At this point the shortened proinsulin chain is transferred from the RER to the Golgi apparatus via secretory vescicles. It is here in the Golgi apparatus that proinsulin fully matures into insulin through the cleavage of the c-peptide by the enzymes Proprotein convertase 2 (PC2) and proprotein convertase 1(PC 1/3) (DeFronzo et al. 2015) **(Figure 23)**.

Insulin Structure

When in its fully matured state, insulin actually exists as a di-peptide monomer formed of the remaining A and B chain. Each of these chains varies slightly in their secondary structure. The A chain, which is 21 residues long, contains an intra di-sulphide bond and an N-terminal helix linked to an anti-parallel C-terminal helix. The B chain on the other hand is 30 residues in length and has a central helical segment. The two chains are linked together via two inter di-sulphide bridges resulting in a protein structure 51 amino acids long and 5802Da in weight (Schmidt et al. 1992).

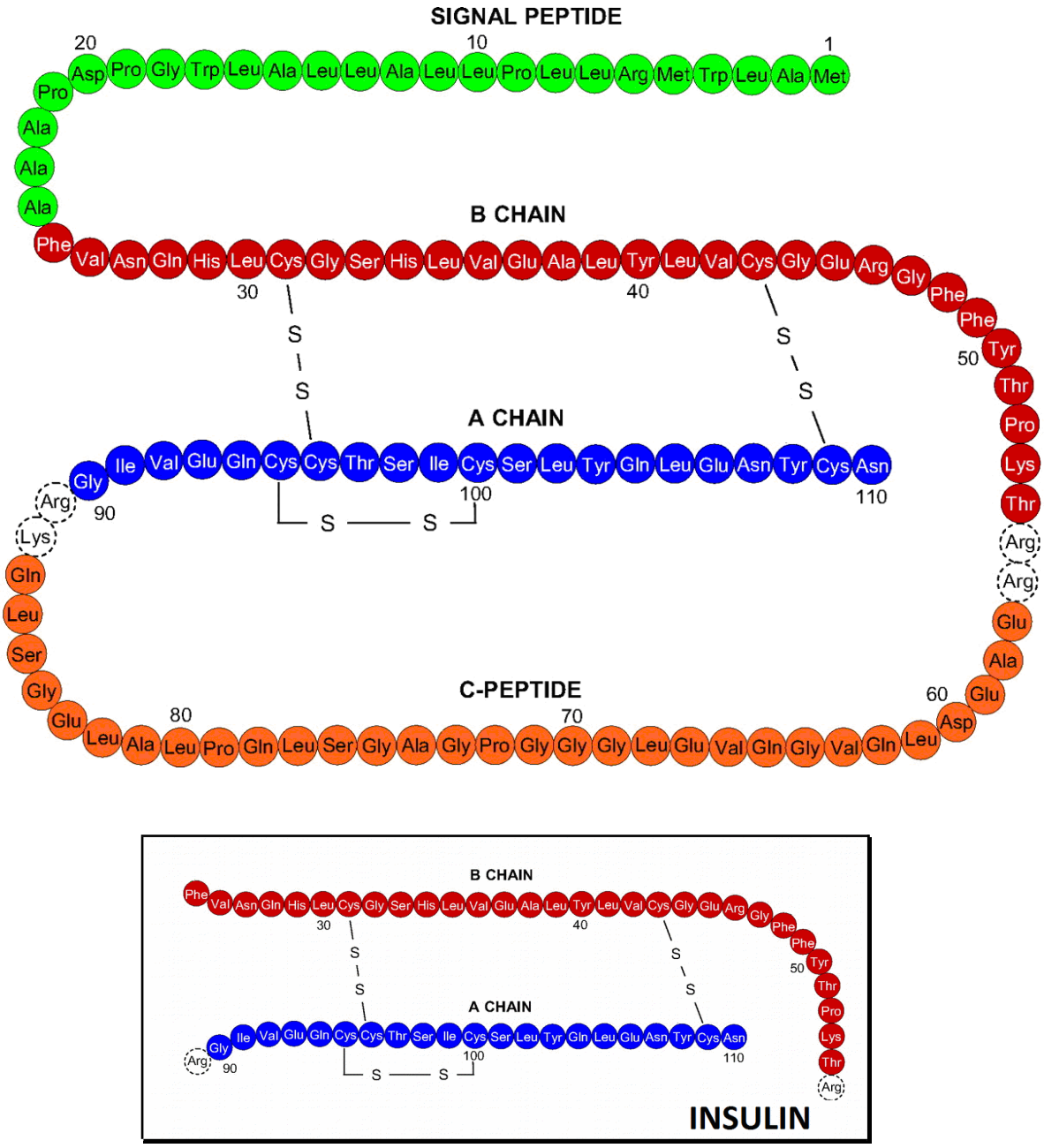


Figure 23: (Above) Human preproinsulin molecule. The amino acid residues which construct the signal peptide are represented in green, the A chain is represented in blue, the B chain is in red and the C chain in orange. In order to produce insulin (below), initially the preproinsulin is translocated from the pancreatic β-cells to the rough endoplasmic reticulum causing the signal peptide to be cleaved in the process. The c-chain is subsequently released through prohormone converstases and carboxypeptidase E cleavage at arginine-65 and arginine -32 (Støy et al. 2007).

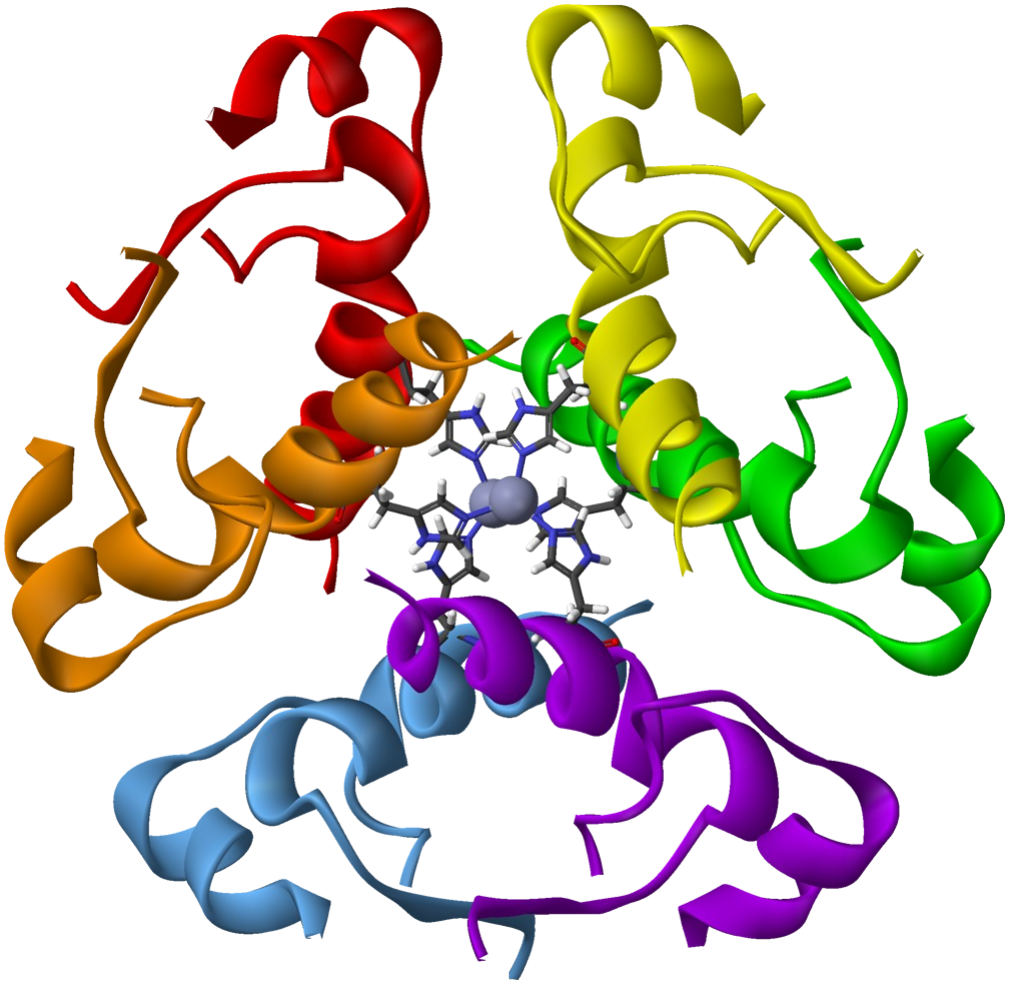


Figure 24: Ribbon diagram of an insulin hexamer. The structure is composed of six insulin monomers which are held together by two zinc ions (depicted as stick figures). Each monomer is represented in a different colour and is composed of 51 amino acids. The monomers each have a molecular mass of 5.8kDa resulting in a hexamer structure approximately 36kDa in size.

At physiological conditions, insulin has the ability to exist as a mixture of monomers, dimers and hexamers. Although it is the monomer which is involved in glucose regulation, the predominant form is actually the 36kDa hexamer. The hexamer acts as a 'storage' structure composed of three dimers and stabilised by two zinc ions. When in the hexameric state, insulin is extremely stable, insoluble and resistant to conformation changes. Due to the increased stability of this state, it is essential to dissociate the hexamers into monomers via removal of the zinc ions in order for fibre formation to occur. It is the resulting monomers which are able to undergo the necessary structural changes for fibre nucleation **(Figure 24)**.

Insulin Derived Amyloidosis

Insulin amyloidosis is rare. It occurs at insulin injection sites of diabetics in the form of a hard nodule (Shikama et al. 2010). Its cause has not been established although some believe it to be due to repeated trauma at the injection site and increased localisation/concentration of the recombinant protein (Lonsdale-Eccles et al. 2009).

Localized insulin amyloid was first reported in 1983 by Störke*l et a*l. but since then, fewer than 100 cases have been described in literature. Due to this lack in experimental data, the exact mechanism of insulin derived amyloidosis is not well understood. That being said, in the year 2014 there were 52 cases of insulin derived amyloidosis reported from one institute alone and seven from another. This would indicate that insulin derived amyloidosis is more common than previously thought and just not being reported (D'Souza et al. 2014) (Nagase et al. 2014).

The underreporting of insulin derived amyloidosis could be due to the misdiagnosis of local site reactions. It is often found that the lack of awareness and availability of histopathological investigations leads to insulin derived amyloidosis being mistaken for lipohypertrophy; a condition similar to insulin amyloidosis which forms fatty lumps on the surface of the skin at

injection sites. Lipohypertrophy is a fairly common side effect of insulin injections if multiple injections are administered to the same part of the skin over a period of time. However, the fatty lumps are not as solid and firm as those seen in amyloidosis. It has also been noted that lipohypertrophy lobules generally regress after cessation of injections, whereas localized amyloidosis does not. It is only after symptoms worsen than a histological examination for an accurate diagnosis is performed instead (Gupta, Singla, and Singla 2015).

In fact, insulin derived amyloidosis can be diagnosed using both histophathology and immunohistochemistry methods. Initially, a sample of extracellular disposition is collected and stained with Congo red. If the apple green birefringence confirms that the substance is amyloid in nature, the sample is subsequently immuno-stained with insulin antibodies to prove that the protein involved is in fact insulin (Gupta, Singla, and Singla 2015).

Symptoms and Treatment

Various symptoms are associated with insulin amyloidosis. The most common being the presence of an amyloid mass at the injection site which is responsible for poor penetration of insulin and probably contributes to insulin resistance. Other symptoms include episodes of severe hypoglycema caused by erratic absorption from the nodule (Endo et al. 2010) and unpredictable releases of insulin causing severe fluctuations in blood sugar levels. It is also common for patients to experience a dramatic gain in weight due to an increased insulin doses prescribed to counteract impaired absorption (Shikama et al. 2010).

Nagase *et al*. compared serum insulin levels before and after injections into the amyloid site. They found that a large portion of the injected insulin did not appear in the serum and suggested that the insulin was either being degraded at the site or trapped. The latter of these seems the more probable explanation when you consider that amyloid fibrils recruit new molecules during growth.

Symptoms of insulin amyloidosis can be reversed by surgically removing the nodule. Constant changing of the injection site can also have a positive effect however both of these solutions are temporary and nodules often reform. As of yet a lasting treatment has not been determined.

Insulin Fibre Formation *in vitro*

The mechanism behind acidic insulin fibre formation *in vitro* has not been fully scrutinised however in 1997 Brange *et al*. published their findings on the matter. They note that the shielding of hydrophobic domains is the main driving force for the partial unfolding of insulin molecules into aggregates. Crystal data presents this in more detail. It shows that the conformational displacement of the B-chain C-terminal with exposure of non-polar, aliphatic core residues including A2, A3, B11 and B15, plays a crucial role in the fibrillation process. Further analysis of this data also suggests that this newly exposed domain interacts with a hydrophobic surface domain formed by the aliphatic residues A13, B6, B14, B17, and B18; a domain normally buried within the hexamer complex.

Whether this is true for all insulin fibres is yet to be seen. As of yet there is not enough published data to either corroborate or dismiss these discoveries. Another point to consider is that the above results ere obtained in acidic conditions. This is not representative of the neutral pH experiments which are to be conducted in this chapter. Therefore the information should only be used as a guideline.

Methods

Insulin Preparation

Bovine insulin powder was obtained from Sigma Aldrich.

In preparation of fibre formation, 20mg of the product was solubilised in 1ml alkaline buffer containing ethylenediaminetetraacetic acid (EDTA) (100mM TRIS, 5mM EDTA, pH10). Immediately following solubilisation, the insulin was dialysed against neutral “TE buffer” (20mM TRIS, 150 NaCl, 0.01% Na Azide, pH8) overnight at a ratio of 1:2000, sample to dialysing buffer respectively, in order to prevent protein deamination from occurring.

It should be noted that an alkaline buffer was chosen in favour of a neutral one due to insulin’s low solubility at neutral pH. Also, it was essential that EDTA be included so to sequester the zinc ions which hold the insulin hexamer together. This in turn would allow dissociation into the monomers which are essential to fibre formation.

Insulin Fibre Formation

A 1 ml sample of 1mg/ml dialysed insulin was prepared in a 1.5ml eppendorf using TE buffer. The sample was continuously inverted - or “flipped” - at a speed of 20rpm on a revolving Stuart Rotator SB3 for 10 days at room temperature and the results were analysed using the ThT fluorescence method described below in ‘Insulin Fibre Analysis with Thioflavin T’.

Example set up:

Sample	Insulin (20mg/ml)	TE Buffer	Final Insulin Concentration
1	50 µl	950 µl	1 mg/ml

This protocol was repeated 3 times. Although an average was not necessary at this point in the investigation, it was important to ensure that fibres could be replicated to similar, if not the same amounts each time. This would allow any future results to be comparable to one another. If each sample has the potential to make the same amount of fibres in the same conditions the experiment is fair. If fibre amounts changed dramatically each time the protocol was carried out it would mean that further into the investigation when SAP was involved, one would not be able to tell if differences in fibre mass were due specifically to the presence of SAP or due to another unknown factor.

Monitoring the Effect of Insulin Concentration on Fibre Formation

To monitor the effect of concentration on insulin fibre formation the above method - “Insulin Fibre Formation” - was repeated using varying concentrations of insulin between 0.1 and 2mg/ml. Each 1ml sample was continuously inverted at a speed of 20rpm on a revolving Stuart Rotator SB3 for 10 days at room temperature and the results were analysed using the ThT fluorescence method described below in ‘Insulin Fibre Analysis with Thioflavin T’.

Example setup:

Sample	Insulin (20mg/ml)	TE Buffer	dH_20	Total Insulin Concentration
2a	100 µl	900 µl	-	2 mg/ml
2b	50 µl	900 µl	50 µl	1mg/ml
2c	25 µl	900 µl	75 µl	0.5mg/ml
2d	20 µl	900 µl	80 µl	0.4 mg/ml
2e	15 µl	900 µl	85 µl	0.3 mg/ml
2f	10 µl	900 µl	90 µl	0.2 mg/ml
2g	5 µl	900 µl	95 µl	0.1 mg/ml

Monitoring the Effect of Speed on Fibre Formation

To monitor the effect of speed on insulin fibre formation the above method "Insulin Fibre Formation" was repeated using varying rotator speeds: 2rpm (minimum speed of rotator), 20rpm or 40rpm (maximum speed). Each 1ml sample was continuously inverted at an individual speed on a revolving Stuart Rotator SB3 for 25 days at room temperature. The results were analysed using the ThT fluorescence method described below in 'Insulin Fibre Analysis with Thioflavin T'.

Example setup:

Sample	Insulin (20mg/ml)	TE Buffer	Rotator Speed
3a	50 µl	950 µl	2 rpm
3b	50 µl	950 µl	20 rpm
3c	50 µl	950 µl	40 rpm

SAP Preparation

Human SAP was prepared by a third party as described previously by Hawkins *et a*l in 1991. In brief, SAP was captured from serum or malignant effusion fluid by its calcium dependent binding to phosphoethanolamine. It was then covalently immobilized by carbodimide coupling to carboxyhexyl sepharose and washed in the presence of calcium. The SAP was subsequently eluted by calcium chelation with EDTA. Trace impurities were removed by reverse affinity chromatography absorption and gel filtration in the absence of calcium which yielded intact, native, fully functional SAP at greater than 99% purity and free of aggregate subunits.

Fibre Formation with SAP Decamers

To observe the effect of SAP decamers on fibre growth, the “Insulin Fibre Formation” method was repeated with the addition of 0.2mg/ml SAP. This provided a stoichiometric ratio of 1:100 of SAP decamers to insulin respectively. Each sample was continuously inverted at a speed of 20rpm on a revolving Stuart Rotator SB3 for 10 days at room temperature. The results were analysed using the ThT fluorescence method described below in ‘Insulin Fibre Analysis with Thioflavin T’. .

Example setup:

Sample	Insulin (20mg/ ml)	TE Buffer	dH_20	SAP (6.66mg/ ml)	Final SAP Concentration
4a	50 µl	920 µl	30 µl	-	-
4b	50 µl	920 µl	-	30 µl	0.2mg/ml

In order to monitor the affect of different SAP decamer concentrations on fibre formation the “Insulin Fibre Formation” method was repeated with the addition of variable SAP decamer concentrations ranging from 0 to 1mg/ml. A slight alteration was made to the protocol in this case. Samples were made up to 0.5ml instead of the normal 1ml in order to conserve insulin stocks. It was not necessary for these samples to be directly comparable to the previous experiments, only to each other. Each sample was continuously inverted at a speed of 20rpm on a revolving Stuart Rotator SB3 for 10 days at room temperature. The results were analysed using the ThT fluorescence method described below:

Insulin Concentration	SAP Decamer Concentration
1 mg/ml	1 mg/ml
1 mg/ml	0.8 mg/ml
1 mg/ml	0.6 mg/ml
1 mg/ml	0.4 mg/ml
1 mg/ml	0.3 mg/ml
1 mg/ml	0.2 mg/ml
1 mg/ml	0.1 mg/ml
1 mg/ml	0.05 mg/ml
1 mg/ml	0

Fibre Formation with SAP Pentamers

To perform this particular assay the SAP was diluted to 0.2mg/ml using TC buffer. This was left to incubate for half an hour to ensure saturation of SAP calcium pockets. Following this, 1mg/ml insulin was added and the sample was continuously inverted at a speed of 20rpm on a revolving Stuart Rotator SB3 for 10 days at room temperature. The results were analysed using the ThT fluorescence method described below in 'Insulin Fibre Analysis with Thioflavin T'.

Example setup:

Sample	Insulin (20mg/ ml)	TC Buffer	dH_20	SAP (6.66mg/ ml)	Final SAP Concentration
5a	50 µl	920 µl	30 µl	-	-
5b	50 µl	920 µl	-	30 µl	0.2mg/ml

In order to monitor the affect of different SAP pentamer concentrations on fibre formation the "Insulin Fibre Formation" method was repeated with the addition of variable SAP pentamer concentrations ranging from 0 to 1mg/ml. A slight alteration was made to the protocol in this case. Samples were made up to 0.5ml instead of the normal 1ml in order to conserve insulin stocks. It was not necessary for these samples to be directly comparable to the previous experiments, only to each other. Each sample was continuously inverted at a speed of 20rpm on a revolving Stuart Rotator SB3 for 10 days at room temperature. The results were analysed using the ThT fluorescence method described below.

Insulin Concentration	SAP Concentration
1 mg/ml	1 mg/ml
1 mg/ml	0.8 mg/ml
1 mg/ml	0.6 mg/ml
1 mg/ml	0.4 mg/ml
1 mg/ml	0.3 mg/ml
1 mg/ml	0.2 mg/ml
1 mg/ml	0.1 mg/ml
1 mg/ml	0.05 mg/ml
1 mg/ml	0

Controls

The high salt concentration required for stabilisation in the TC buffer was a variable that had to be monitored. To see the effect the salt was having on insulin fibre formation the assay had to be repeated in a High Salt buffer ("THS Buffer" - 20mM TRIS, 600mM NaCl, 0.01% Na Azide, pH8). 1mg/ml of insulin was continuously inverted at a speed of 20rpm on a revolving Stuart Rotator SB3 for 10 days at room temperature in the presence/absence of 0.2mg SAP.

The THS buffer was identical to the TC buffer but lacked the calcium chloride. If fibre behaviour in the THS assay matched that of the original TE buffer, it could be presumed that the additional salt had no effect on fibre formation. This would therefore suggest that all differences noted in the TC buffer assay were due to the change in SAP conformation (or the calcium) and not the salt. If the THS assay did not match that of the original TE assay, one would have to presume that the salt does indeed have an affect on fibre formation. This would mean that the conditions for SAP pentamer interactions would have to be reviewed and altered.

Example setup:

Sample	Insulin (20mg/ml)	THS Buffer	SAP (6.66mg/ml)	dH_20
6a	50 µl	920 µl	30 µl	-
6b	50 µl	950 µl	-	30 µl

It was also important to ensure that any fibres formed during the course of these experiments were made purely of insulin. To test whether SAP has the

ability to form fibres under these conditions 0.2mg/ml of SAP was rotated on a revolving Stuart Rotator SB3 for 50 days at room temperature.

Sample	SAP (6.66mg/ ml)	TE Buffer	TC Buffer	THS Buffer
Control a	30 µl	920 µl	-	-
Control b	30 µl	-	920 µl	-
Control c	30 µl	-	-	920 µl

Fibre Analysis with Thioflavin T

Thioflavin T (ThT) was used in order to confirm and quantify fibre growth. ThT is a benzothiazole dye which, upon binding amyloid, increases in fluorescence at a wavelength of 420nm. It binds proportionally to the channels which run parallel to the β-sheet making it a useful quantitative method of analysis.

In preparation for analysis, a 9mM stock solution of ThT was made, wrapped in foil and refrigerated for up to two months. On the day of measurement, the stock solution was diluted to 500 ml of a concentration of 0.1µM. Fibre samples were shaken vigorously on a vortex mixer for 15 seconds and a 5µl sample was taken and added to the diluted ThT. The mixture was incubated at room temperature, in the dark for 15 minutes and then measured on a Perkin Elmer LS 55 Luminescence Spectrometer (excitation wavelength 442 nm, slit width 5nm). Please note that buffer containing ThT was used as a base line in these measurements and subtracted from each result.

Fibre Detection using Congo Red

To perform the Congo red analysis, 3 layers of 50µl samples were individually dried onto a glass slide and stained using Mayer's Hematoxylin. Following this, the slides were placed in 0.34M alcoholic sodium chloride solution for 20 minutes, drained and replaced with 3mM Congo red solution for a further 20 minutes. Finally the slides were dehydrated using ethanol and cleared with xylene. The slides were then viewed under a microscope using polarised light.

Structural Analysis using Transmission Electron Microscopy (TEM)

In preparation of TEM analysis, copper grids were covered in carbon coated formvar resin. Samples 3µl in size were transferred onto the grids and following three minutes of incubation, had the excess liquid absorbed using a torn edge of hardened, ash-less filter paper. A 3µl measurement of 2% uranyl acetate solution was used to stain the sample and a further three minutes of incubation was carried out. Again, the excess solution was absorbed using the filter paper and the samples were left to air dry. The slides were then examined under the electron microscope.

Structural Analysis using Circular Dichroism

The samples chosen to undergo CD analysis were as follows:

Sample	Insulin (20mg/ml)	Buffer Type	Buffer	dH_20	SAP (6.66mg/ml)	Final SAP Concentration
4a	50 µl	TE	920 µl	30 µl	-	-
4b	50 µl	TE	920 µl	-	30 µl	0.2mg/ml
5a	50 µl	TC	920 µl	30 µl	-	-
5b	50 µl	TC	920 µl	-	30 µl	0.2mg/ml
6a	50 µl	THS	920 µl	30 µl	-	-
6b	50 µl	THS	950 µl	-	30 µl	0.2mg/ml

Sample Preparation

In order to perform CD all the samples had to be transferred to water. This was due to the fact that the calcium chloride present in the TC buffer solutions shows high absorbance in the far UV region and often disrupts CD results (Kelly, Jess, and Price 2005). Water is a non chiral solution which would not affect the results. Transference of each sample to water was achieved using the following protocol:

Each fibre sample was spun in a microcentrifuge for 6 minutes at 13,000rpm so to separate the solid fibre mass from the solution. Following this, the supernatant was drained from each and replaced with water. The fibre masses were “washed” via vigorous vortexing until re-suspended. The entire process was repeated twice more to ensure complete removal of buffer, in particular, those containing calcium chloride.

In the cases where SAP was present, 10mM ethylene glycol tetraacetic acid (EGTA) was also added to the first wash so to chelate the calcium and thus unbind SAP from the fibres. The EGTA was later removed with subsequent washes so to prevent interference with the results.

Circular Dichroism

Once suspended in water the samples needed to be made to the same concentration so to allow comparative analysis. Unfortunately, due to the fact that the protein to be measured via CD was in solid form, this was deemed impractical. Solid protein concentrations can only be measured accurately using chemical analysis such as bicinchoninic assays; a time consuming process. Optical densities, although not the most precise method for measuring protein concentration, are much quicker and more practical to determine so were used instead.

Each sample was diluted to the same optical density through the addition of more or less distilled water. Following this, they were individually loaded into a custom built rotating cell. A rotating cell was chosen instead of a stationary one in an attempt to keep the fibres in suspension during data collection. The difference of absorption was measured between wavelengths of 180-260 at a slit size of 0.5mm and a path length of 0.2mm. Distilled water was used as a base line and subtracted from each result.

Controls

As controls, 0.5mg/ml native insulin, 0.2mg/ml native SAP and SAP aggregate were used. The aggregated SAP was created through the addition of a few drops of 37% hydrochloric acid to a 0.2mg/ml sample of SAP. Once the precipitate had formed, it was transferred into distilled water via centrifugation, removal of the supernatant and replacement with water. The new contents were mixed and the diluted until the OD was identical to that of the samples.

Bicinchoninic Acid Assay

BCA is a method used to determine the protein concentration of a solution between the concentrations of 0.5 µg/ml and 1.5 mg/ml. It incorporates all protein present including solubilised protein and aggregated protein whether present as fibres or amorphous aggregate. By removing the solubilised protein from each sample, a comparison between the BCA and ThT results could be made.

BCA is a highly alkali solution (pH 11.25) which contains the following chemicals: bicinchoninic acid, sodium carbonate, sodium bicarbonate, sodium tartrate and copper (II) sulphate pentahydrate. It exhibits a colour change of green to purple proportional to protein concentration. This associated colour change is a result of two chemical reactions. The first is named the biuret reaction and involves the reduction of the Cu^{2+} ions from the copper (II) sulphate pentahydrate into Cu^{+}. This is done by the cysteine, tyrosine and tryptophan side chains present in the protein and is a temperature dependent reaction. The result of this particular reaction occurring is the formation of a blue/green colour. The second reaction occurs when two bicinchoninic acid molecules chelate with these newly formed Cu^{+} ions. A purple product is produced which absorbs at wavelength 562nm. The more cysteine, tyrosine and tryptophan present the more copper is reduced. The higher the concentration of reduced copper the higher the amount of chelation that

occurs. And the more chelation that occurs the deeper the purple colour becomes, hence the BCA reaction being proportional to protein concentration. ("Chemistry of Protein Assays" 2015) **(Figure 25)**.

The following samples were submitted to BCA analysis:

Sample	Insulin (20mg/ml)	TE Buffer	dH_20	SAP (6.66mg/ml)	Final SAP Concentration
4a	50 µl	920 µl	30 µl	-	-
4b	50 µl	920 µl	-	30 µl	0.2mg/ml
5a	50 µl	920 µl	30 µl	-	-
5b	50 µl	920 µl	-	30 µl	0.2mg/ml
6a	50 µl	920 µl	30 µl	-	-
6b	50 µl	950 µl	-	30 µl	0.2mg/ml

In order to prepare the samples for BCA analysis, they were each spun on a microcentrifuge for 10 minutes to separate the insoluble protein aggregate form the solubilised solution. The supernatant containing soluble insulin was discarded and replaced with distilled water. Each sample was vigorously mixed on a vortex until resuspended.

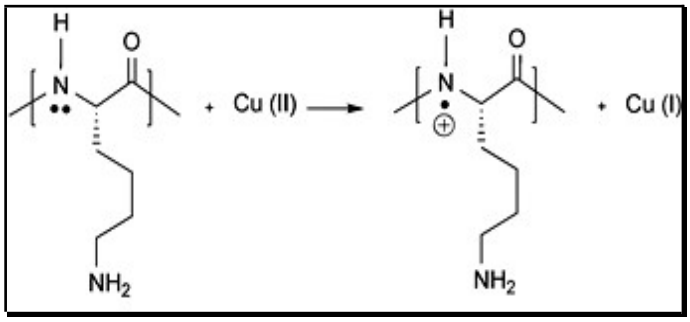
To perform the bicinchoninic acid (BCA) assay, 0.5ml of a 0.74M solution of copper (II) chloride was made and mixed with 2.5ml of BCA reagent. At this point a green colour was observed. A 1:20 ratio of protein sample to BCA solution was made and incubated at 37°C for 30 minutes. These were allowed to cool to room temperature and the resulting coloured solution was measured on a NanoDrop 3300 fluorospectrometer at 562nm. The NanoDrop results were converted to concentrations through use of a calibration curve and the results were compared to their ThT equivalents.

The percentages of native, amorphous and cross β-sheet protein were calculated as follows: The native protein concentrations were measured from the removed supernatant on a NanoDrop 3300 fluorospectrometer at a wavelength of 280nm. These figures were compared to the initial mass of insulin used and converted into a percentage. The amorphous and cross β-sheet proteins were calculated via comparisons of the ThT and BCA data of

each sample. A calibration curve was made using 12 proteins of known concentrations. The figures derived from this graph were then compared to the initial mass of insulin used and converted into a percentage.

Fibre disassembly using SAP Decamers

Dialysed insulin was diluted using TE buffer to make 0.6ml of a 1mg/ml solution and fibres were formed using the above method (Insulin Fibrillogenesis). After ten days of rotation at room temperature on a revolving Stuart Rotator SB3 (20rpm) the fibres were removed, thoroughly mixed using a vortex and had 0.1mg of SAP decamers added to the eppendorf. The eppendorf was rotated at a slower speed of 2rpm so to keep all components in suspension for a total of 40 days. Additional aliquots of 0.1mg SAP were added every 10 days and fibre amounts were measured daily via ThT measurements on a spectrophotometer as in the “Fibre analysis with Thioflavin T” section above.



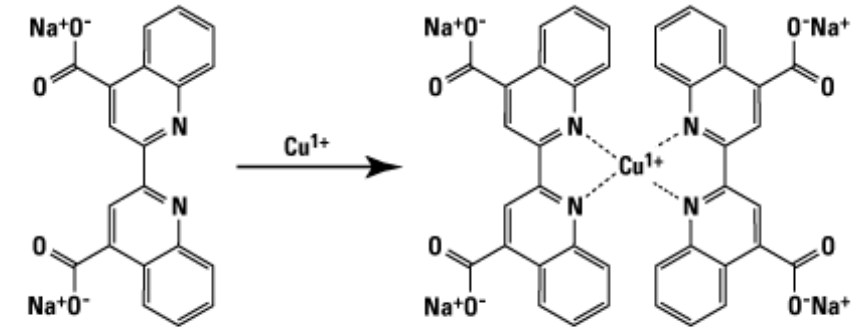


Figure 25: (Above) Diagram showing the biuret reaction –the first step in a BCA reaction. It involves the reduction of the Cu^{2+} ions from the copper (II) sulphate pentahydrate in the BCA solution into Cu^{+} by the cysteine, tyrosine and tryptophan side chains present. This is a temperature dependent reaction. (Below) The second reaction involves the chelation of a Cu^{+} ion by two bicinchoninic acid molecules creating a purple product which absorbs at wavelength 562nm. The more peptide bonds present the more copper reduced (Chemistry of Protein Assays, 2015).

Results and Discussion

Insulin Fibre Analysis with ThT

A sample of dialysed insulin was diluted to 1mg/ml in TE buffer (pH 8) and continuously inverted or 'flipped' on a rotator at a speed of 20rpm in a 1.5ml eppendorf. This was done at room temperature for a total of 10 days.

Within the first 2 days of inversion - or 'flipping' - a white precipitate developed. This precipitation increased in density over the course of 3-6 days and then appeared to remain constant for the rest of the protocol. On the tenth day the newly precipitated insulin sample underwent ThT analysis in order to determine whether the precipitate was cross β-sheet in nature or amorphous aggregate. The ThT results show enhanced fluorescence between wavelengths 420nm and 600nm with a peak of 652 at 476.5 nm. This provides a relatively good indication that the precipitate is cross β-sheet in nature, however, this is not known for certain as ThT can often give false positive results **(Figure 26)**.

The 10 day 'rotator method' was repeated using two more samples of 1mg/ml dialysed insulin and ThT analysis was carried out. This was not to obtain an average as an average was not necessary at this point. Instead, the repetition of this assay was to ensure that fibres could be replicated to similar, if not identical amounts each time. The ability to replicate fibres to a similar mass or density is vital as it allows future results to be comparable to one another. If each sample has the potential to make the same amount of fibres in the same conditions the experiment is fair. If fibre amounts change dramatically each time it would mean that further into the investigation, when SAP is involved, one will not be able to tell if differences in fibre mass were due specifically to the presence of SAP or due to another unknown factor.

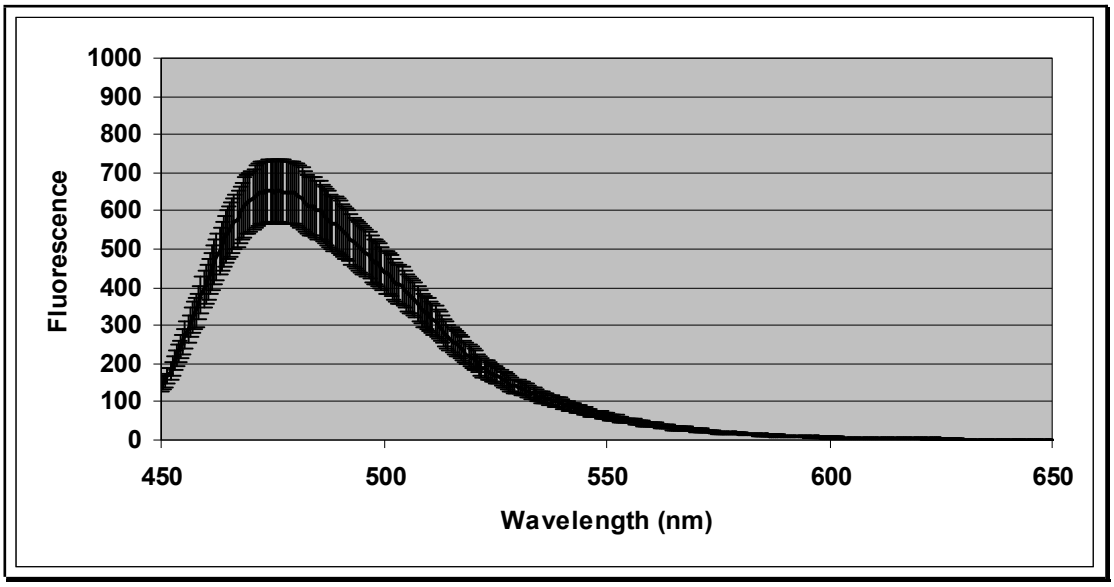


Figure 26: Graph illustrating ThT fluorescence of insulin fibres made in TE buffer as measured on a Perkin Elmer LS 55 Luminescence Spectrometer (excitation wavelength 442 nm, slit width 5nm). Fibres were rotated for 10 days on a revolving Stuart Rotator SB3 at 20rpm at room temperature. Fluorescence intensity peaks at 652 at a wavelength of 476.5 nm and instrumental errors are plotted.

Upon analysis of these two extra samples it was found that fluorescence intensity is not as consistent as expected. The ThT measurements of the subsequent samples are similar to one another but quite different to the original sample. They exhibit peaks in fluorescence at 988 and 998 thus providing a range of approximately 600-1000 for all three values. This was unexpected as each had undergone identical conditions **(Figure 28)**.

The fluctuations seen between samples 2 and 3 - which obtained florescence readings of 988 and 998 - are probably a result of the dynamic nature of cross β-sheet fibres. If fibres are constantly detaching and reforming themselves, at any one given time one may experience more or less fluorescence compared to another. The larger difference seen in the first sample (ThT reading 652) however cannot be explained that way. Instead, it is most likely that this particular reading is an anomaly. This sample was the first attempt of a new protocol so could quite possibly be subject to handling errors.

The effect of rotator speeds and insulin concentration on cross β-sheet fibre formation were tested. The results were as follows:

Changes in rotator speeds had no effect on total fibre amounts according to the ThT readings. Figure 27 shows each sample reaching a peak of approximately 1000 by the end of the 25 day protocol. The different rotator speeds, however, did have an impact on growth rates. The assays performed at slower rotator speeds had an extended lag period compared to those at faster speeds. Fibres grown in 2rpm conditions experienced a lag phase of approximately 10 days where as fibres grown in 20 and 40rpm experienced shorter lag phases of 5 and 3 days respectively. This is logical as the faster the samples are being flipped/rotated, the higher the continuous movement of surface tension and the higher the internal pressure of the sample during rotation. This provides more energy to break the very high surface tension value given by water and thus creates more force to denature the insulin monomers.

Insulin concentration was also seen to have an affect on total fibre mass. The more protein that was present in a sample the higher the fluorescence intensity and vice versa. The two components are directly proportional to one another and experience a strong positive correlation of 0.897 **(Figure 32)**.

Alternate Fibre analysis

As previously mentioned, ThT is often considered unreliable as a diagnostic tool as it has the propensity to bind bacteria and other tissue matter. Although it is highly unlikely that any tissue matter would have been present in these particular assays as samples were not extracted from *ex vivo* specimens, it is entirely possible that bacteria could have contaminated them. Therefore, in order to confirm the presence of cross β-sheet fibres the samples were submit to alternative analytical approaches including Congo red analysis and TEM.

When submit to Congo red analysis, all samples clearly demonstrated the apple green birefringence typical of cross β-sheets. This alongside the ThT results provides a strong indication that the insulin had indeed formed cross β-sheet fibres and not amorphous aggregate. Nevertheless, Congo red analysis can often be as unreliable as ThT as a diagnostic tool but for a completely different reason. Unlike ThT, Congo red is not structure specific and so can be seen to bind to native proteins as well as their fibre counterparts. Although this has not proven to be true of all native proteins, it adds an element of uncertainty to any results obtained by this method. Due to this, it was still necessary to perform TEM analysis. Of which, the resulting images also confirmed the presence of cross β-sheets. At a magnification of x140,000, rope-like structures approximately 7nm in diameter and of indefinite length are clearly visible.

In summary, unlike previous studies which produced insulin fibres only in acidic conditions, the above method produced large amounts of cross β-sheet insulin fibres in a pH very close to physiological conditions. This demonstrates that insulin fibres can indeed be made *in vitro* at neutral pH. This was proved using ThT, Congo red and TEM analysis.

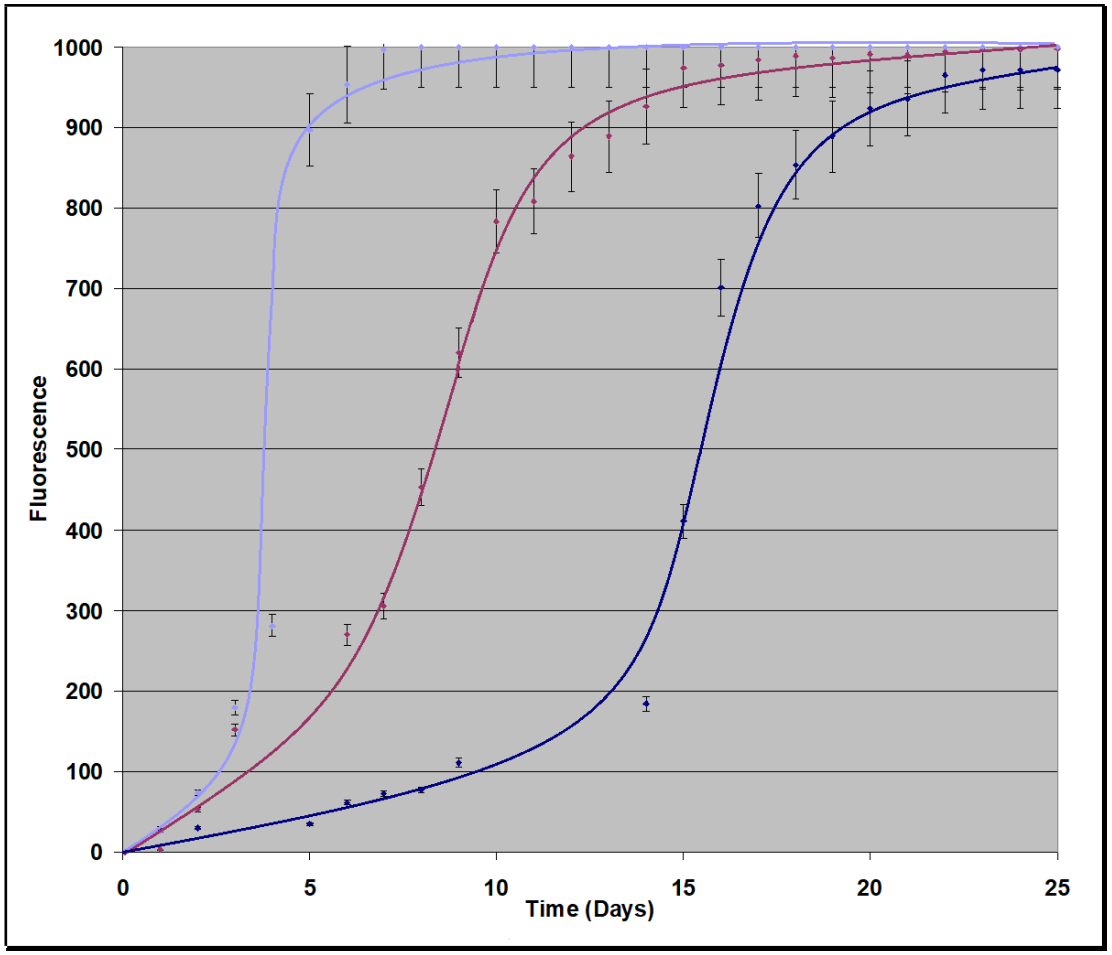


Figure 27:A graph showing the progression of insulin fibre growth in TE buffer. Fibres were rotated at room temperature for 25 days on a revolving Stuart Rotator SB3 set at various speeds: 40rpm (light blue), 20rpm (magenta) and 2rpm (dark blue). The ThT fluorescence indicates that rotator speeds have an impact on growth rates but not on total fibre amounts as each sample reaches a peak of approximately 1000 by the end of the 25 day protocol. This graph exhibits the following pattern: the slower the rotator speed the longer the lag phase of fibre growth. Fibres grown in 2rpm conditions experience a lag phase of approximately 10 days where as fibres grown in 20 and 40rpm experience much shorter lag phases of 5 and 3 days respectively.

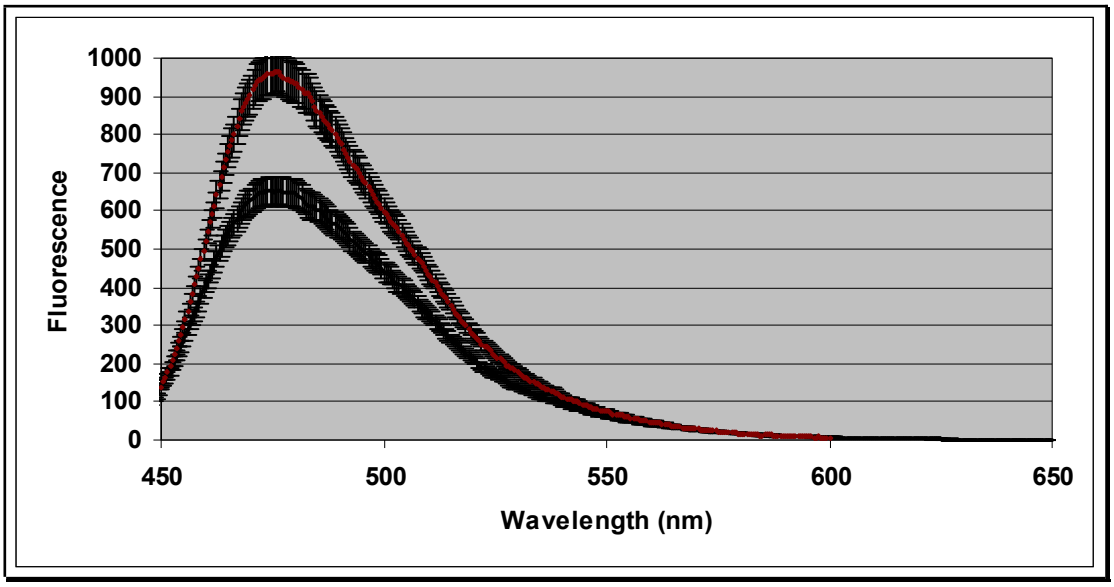


Figure 28: A graph comparing ThT fluorescence of the original insulin fibres made in TE buffer (black) with the two subsequent repetitions of the rotator protocol (maroon). All measurement were obtained on a Perkin Elmer LS 55 Luminescence Spectrometer (excitation wavelength 442 nm, slit width 5nm) Fibres were rotated for 10 days on a revolving Stuart Rotator SB3 at 20rpm at room temperature. The original fibres peak at a fluorescence intensity of 652 at a wavelength of 476.5 nm. The subsequent samples exhibit peaks in fluorescence at 988 and 998. All instrumental errors are plotted.

The Effect of SAP Decamers on Fibre Growth

In order to monitor the effect of SAP decamers on fibre formation the “Insulin Fibre Formation” method was repeated with the inclusion of SAP. Dialysed insulin was diluted to 1mg/ml in TE buffer and continuously inverted or ‘flipped’ on a rotator at a speed of 20rpm in the presence of 0.2mg/ml SAP decamers. Following 10 days of inversion via the rotator method, the resulting white precipitate formed was submitted to ThT fluorescence **(Figure 29)**.

The ThT results show that the fibres grown in the presence of SAP decamers at a substoichiometric ratio of 1:100 (SAP: insulin) demonstrate an average *decrease* in ThT fluorescence of 54.0%. This figure was calculated from an average of three separate assays, each of which underwent Congo red and TEM analysis so to confirm that the precipitate formed was cross β-sheet in nature. It indicates that SAP, in its decameric form, has the capability of inhibiting fibre growth. Whether it achieves this through the depletion of existing fibres or through the prevention of initial fibre formation is unclear at this point.

Again, discrepancies in ThT fluorescence are noted between each sample. The samples range between a fluorescence of 628.1 and 997.8 in the absence of SAP and between a range of 291.8 and 624.8 in the presence of the SAP decamers. This provides an error margin of +14.54%, -27.91% and +35.65%, -36.58% respectively. More specifically, sample 1 of this particular experimental condition has an average decrease in fibres of 36.7% and experiments 2 and 3 have 26.1% and 70.8% respectively. This is quite a high percentage discrepancy; however, each corresponding pair of assays (TE fibres with and without SAP) each experienced a decrease in fibre amounts thus validating these conclusions. Furthermore, in the case of the SAP decamer assays, it is important to remember that the SAP molecule exists as a mixture of both pentamers and decamers. The ratio between the two would be prone to fluctuate between samples and thus explain this large range in percentages.

Insulin 1mg/ml	SAP	Fluorescence /475nm		
		Experiment 1	Experiment 2	Experiment 3
TE Buffer	0	987.6	628.1	997.8
TE Buffer	0.2mg/ ml	624.8	464.4	291.8
% change in fluorescence	-	-36.7	-26.1	-70.8
Average % change	-		**-44.5**	

Table 2: A table showing the effect of SAP on insulin fibrillogenesis made in TE buffer. Values are taken from the peak of ThT fluorescence (475nm). The amount of fibres decreased with the addition of SAP decamers: there was an average of 54% reduction in ThT fluorescence between assays with and without SAP in TE buffer. 1mg/ml insulin was used for each sample and 0.2mg/ml SAP in the designated samples. The samples range between a fluorescence of 628.1 and 997.8 in the absence of SAP and between a range of 291.8 and 624.8 in the presence of the SAP decamers. This provides an error margin of +14.54%, -27.91% and +35.65%, -36.58% respectively. More specifically, sample 1 of this particular experimental condition has an average decrease in fibres of 36.7% and experiments 2 and 3 have 26.1% and 70.8% respectively.

When the above assays were repeated using various SAP concentrations (0.1-1mg/ml) the following results were observed: The samples grown with higher SAP concentrations experienced a larger decrease in total fibre amount whereas those with smaller concentrations experienced a smaller decrease in fibre amounts. The two factors are directly proportional and exhibit a strong negative correlation of -0.812 **(Figure 33)**. These results clearly show that SAPs chaperoning activity is proportional to its quantity. The more SAP molecules present the more SAP:intermediate interactions that are occurring and the fewer fibres that are being produced.

The chaperoning activity of SAP decamers was further confirmed using Congo red analysis. In Figure 35 images A and B represent insulin fibres grown in TE buffer with and without SAP decamers. As expected, image B –

those grown in the presence of SAP decamers – demonstrate a decrease in fibre density compared to those grown without. The decrease in fibres, however, cannot be given a value in this case as the Congo red technique is not a quantitative method.

The mechanism on how SAP decreases fibre amounts cannot be deciphered at this point although theories can be suggested. Perhaps SAP is detaching intermediates from the fibre structure and binding them either reversibly or irreversibly. This seems unlikely as crystal analysis has shown that it is the B-face of SAP which binds to fibres via a calcium mediated interaction, not the A face. Not only is calcium absent in this particular assay but the B-face is unavailable due to the interactions required to make a decamer structure. It is more plausible that SAP is interacting with the intermediates before they become part of the fibre structure or after they, themselves, detach from it. Again the nature of the binding is unknown at this point and can be either reversible or irreversible.

The Effect of SAP Pentamers on Fibre Growth

In order to see the effect of SAP pentamers, the fibrillisation assay had to be repeated in an alternate buffer which contained calcium: TC buffer. Although it is essential to add calcium in order to ensure all SAP molecules exist in their pentameric form, by adding calcium to the assay there is an increased risk of SAP auto-aggregation due to the carboxylate group of glutamic acid 87 of one SAP molecule binding to the calcium dependent ligand pocket of another. As a way to prevent autoaggregation in the calcium buffer, supra physiological concentrations of salt were also included as explained in chapter 1.

A sample of dialysed insulin was diluted to 1mg/ml in TC buffer and continuously inverted or ‘flipped’ on a rotator at a speed of 20rpm in the presence of 0.2mg/ml SAP pentamers. Following 10 days of inversion via the

rotator method, a white precipitate was formed in each sample submitted to ThT fluorescence.

It was found that fibres grown in the presence of SAP pentamers at a substoichiometric ratio of 1:100 (SAP: insulin) have an average 50.7% increase of ThT fluorescence. This is directly opposite to what happened in the presence of SAP decamers. These results indicate that the pentamers act as fibrillogenic enhancers. This function could be achieved through either catalytic activity on fibre growth or through stabilisation of the final structure. The existing literature indicates that the latter is the more plausible explanation as SAP is known to stabilise fibres upon binding. Nonetheless, further testing is required before the former can be dismissed.

As expected, discrepancies in ThT fluorescence are noted between samples of the same condition, however, each corresponding pair of assays (TC fibres with and without SAP pentamers) each experienced an increase in fibre amounts thus validating these conclusions. The samples range between a fluorescence of 194.2 and 311.3 in the absence of SAP and between a range of 269.6 and 515.1 in the presence of the SAP pentamers. This provides an error margin of +17.08%, -26.97% and +27.00 %, -33.93% respectively. Specifically, experiment 1 had an average increase in fibres of 38.8% and experiments 2 and 3 had increases of 65.5% and 48.0% respectively.

		Fluorescence /475nm		
Insulin 1mg/ml	**SAP**	**Experiment 1**	**Experiment 2**	**Experiment 3**

TC Buffer	0 0.2mg/	194.2	311.3	292.1
TC Buffer	ml	269.6	515.1	432.2
% change in fluorescence	**-**	38.8	65.5	48.0
Average % change	**-**		**50.8**	

Table 3:: A table showing the effect of SAP on insulin fibrillogenesis. Values are taken from the peak of ThT fluorescence (475nm). 1mg/ml insulin was used for each sample and 0.2mg/ml SAP in the designated samples. The samples range between a fluorescence of 194.2 and 311.3 in the absence of SAP and between a range of 269.6 and 515.1 in the presence of the SAP pentamers. This provides an error margin of +17.08%, -26.97% and +27.00 %, -33.93% respectively. Specifically, experiment 1 had an average increase in fibres of 38.8% and experiments 2 and 3 had increases of 65.5% and 48.0% respectively.

When the above assays were repeated using various SAP concentrations (0.1-1mg/ml) the following results were observed: The samples grown with higher SAP concentrations experienced a larger increase in total ThT fluorescence whereas those with smaller concentrations experienced a smaller increase in ThT fluorescence. The two factors are directly proportional and exhibit a strong positive correlation of 0.713. These results clearly show that SAPs enhancing activity is proportional to its quantity. The more SAP molecules present the more SAP:intermediate interactions that are occurring and the fewer fibres that are being produced **(Figure 30, Figure 34)**.

When compared to the fibres grown in TE buffer, it is noted that ThT fluorescence in these TC assays are not as high as previously seen. The maximum fluorescence intensity reached for TC fibres is 311 (in the absence of SAP) which is only a third of what the TE fibres are able to achieve. Equal

amounts of aggregate were found in both the TE and TC samples and the same ThT stock was used for all. This implies that the different buffer conditions are having an impact on the ThT results.

ThT fluorescence is considered proportional for a certain type of amyloid fibre but only under uniform conditions. By adding calcium and salt to the buffer there is an increased risk of inhibiting ThT:fibre binding and thus diminishing fluorescence. That being said, when subject to ThT analysis, these fibre samples were diluted by a factor of x100 meaning the calcium and salt levels would be very low overall and not likely to have affected fluorescence quite so dramatically . It is more reasonable to assume that the fault lay within the fibres themselves or lack there of. Potentially the calcium present in the buffer has caused dimerisation of the insulin monomers and thus made them incapable of forming fibres. It is also entirely possible that the morphologies of the TE and TC fibres are quite different having grown in contrasting conditions. This is potentially a factor which would lead to either unfair ThT binding or even the formation of less stable fibres.

Controls

In order to clarify the reason behind the reduced ThT fluorescence in TC fibres, a control assay was set up. Insulin fibres were produced in a high salt buffer - THS buffer. THS buffer was identical to TC buffer but lacked the calcium chloride, thus removing one of the two the additional variables. If ThT fluorescence in the THS assay matched those of the original TE buffer, it could be concluded that the additional salt had no effect on fibre formation and that any differences seen were due to the calcium alone. If ThT fluorescence of the THS assay did not match that of the original TE assay, one would have to presume that the salt did indeed have an affect on fibre formation. This would mean that the conditions for SAP pentamer interactions would have to be reviewed and altered.
ThT fluorescence of the THS buffer was capable of reaching similar peaks as in the original TE buffer. The samples range between a fluorescence of 995.2

and 767.5 giving an error margin of +9.48%, -15.57%. This demonstrates that the increased salt amounts have no adverse affects on fibre growth or ThT fluorescence. When the control THS fibres are grown in the presence of SAP they experience an average decrease of 38.7% in fibre amounts with a range of 377.5 and 619.1 between the three different attempts. This results in an error margin of +17.65 and -28.25%. This is similar to the original TE assays and confirms that in the THS buffer, SAP maintains its decameric structure and function **(Figure 31)**.

		Fluorescence /475nm		
Insulin 1mg/ml	**SAP**	**Experiment 1**	**Experiment 2**	**Experiment 3**
THS Buffer	0	995.2	767.5	964.3
THS Buffer	0.2mg/ml	619.1	377.5	582.1
% change in fluorescence	-	37.8	50.8	39.6
Average % change	-		**42.7**	

Table 4: A table showing the effect of SAP on insulin fibrillogenesis made in THS buffer. Values are taken from the peak of ThT fluorescence (475nm). The amount of fibres decreased with the addition of SAP decamers: there was an average of 42.7% reduction in ThT fluorescence between assays with and without SAP in THS buffer. 1mg/ml insulin was used for each sample and 0.2mg/ml SAP in the designated samples. The samples range between a fluorescence of 767.5 and 995.2 in the absence of SAP and between a range of 377.5 and 619.1 in the presence of the SAP decamers. This provides an error margin of +9.48%, -15.57% and +17.65 and -28.25% respectively. More specifically, sample 1 of this particular experimental condition has an average decrease in fibre amounts of 37.8% and experiments 2 and 3 have 50.8% and 39.6% respectively.

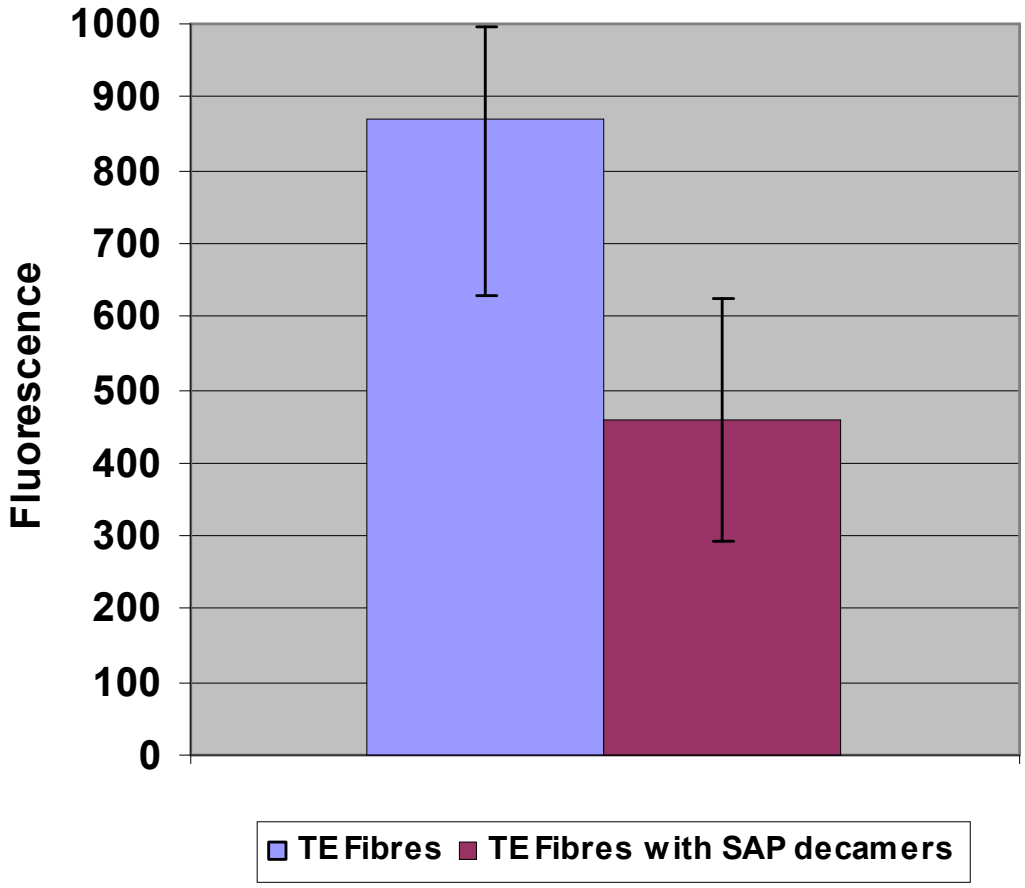


Figure 29: A bar graph comparing average ThT fluorescence of TE insulin fibres grown in the presence/absence of 0.2mg/ml SAP decamers. Fibres were rotated for 10 days on a revolving Stuart Rotator SB3 at 20rpm at room temperature. Each bar is an average of three samples and has error margins included so to show the range of those three samples. The fibres grown in TE buffer produce an average fluorescence of 657.2 where as those grown in the presence of SAP decamers produce an average fluorescence of 292.0. The presence of SAP decamers causes an average reduction of 55.6% in fibre formation.

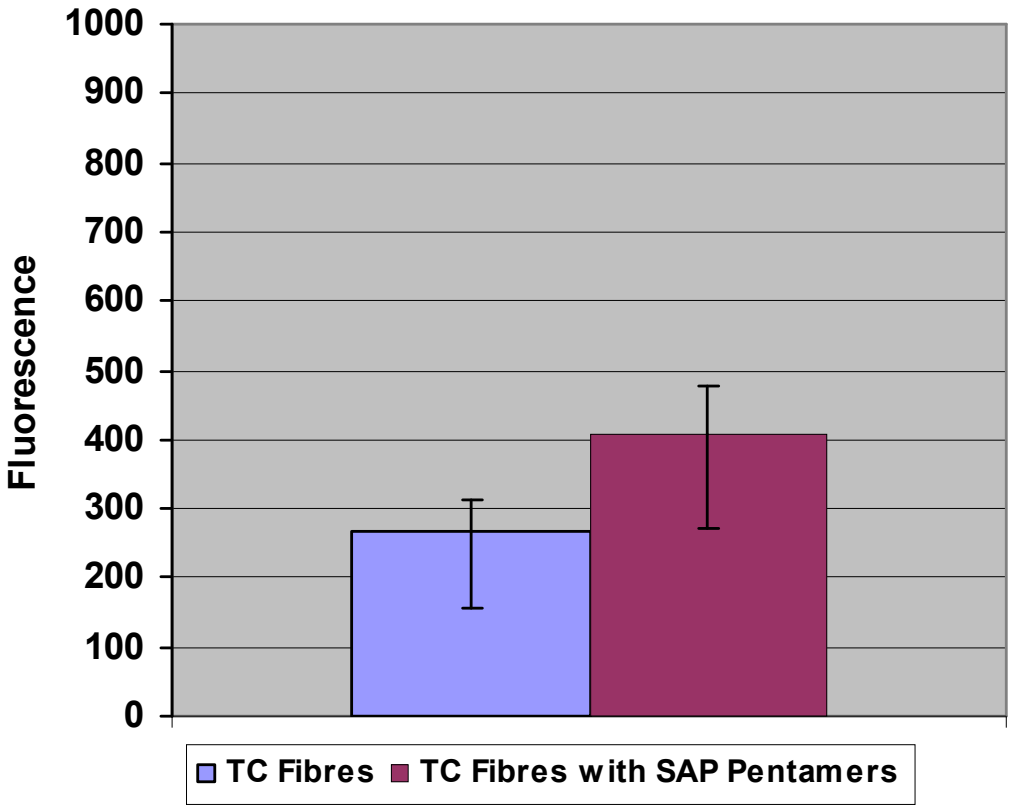


Figure 30: A bar graph comparing average ThT fluorescence of insulin fibres grown in the presence/absence of 0.2mg/ml SAP pentamers in calcium (TC) buffer. Fibres were rotated for 10 days on a revolving Stuart Rotator SB3 at 20rpm at room temperature. Each bar is an average of three samples. Error margins are included to show the range of those three samples. The fibres grown in TC buffer produce an average fluorescence of 265.9 where as those grown in the presence of SAP decamers produce an average fluorescence of 405.3. The presence of SAP decamers causes an average increase in fibre formation of 52.5%.

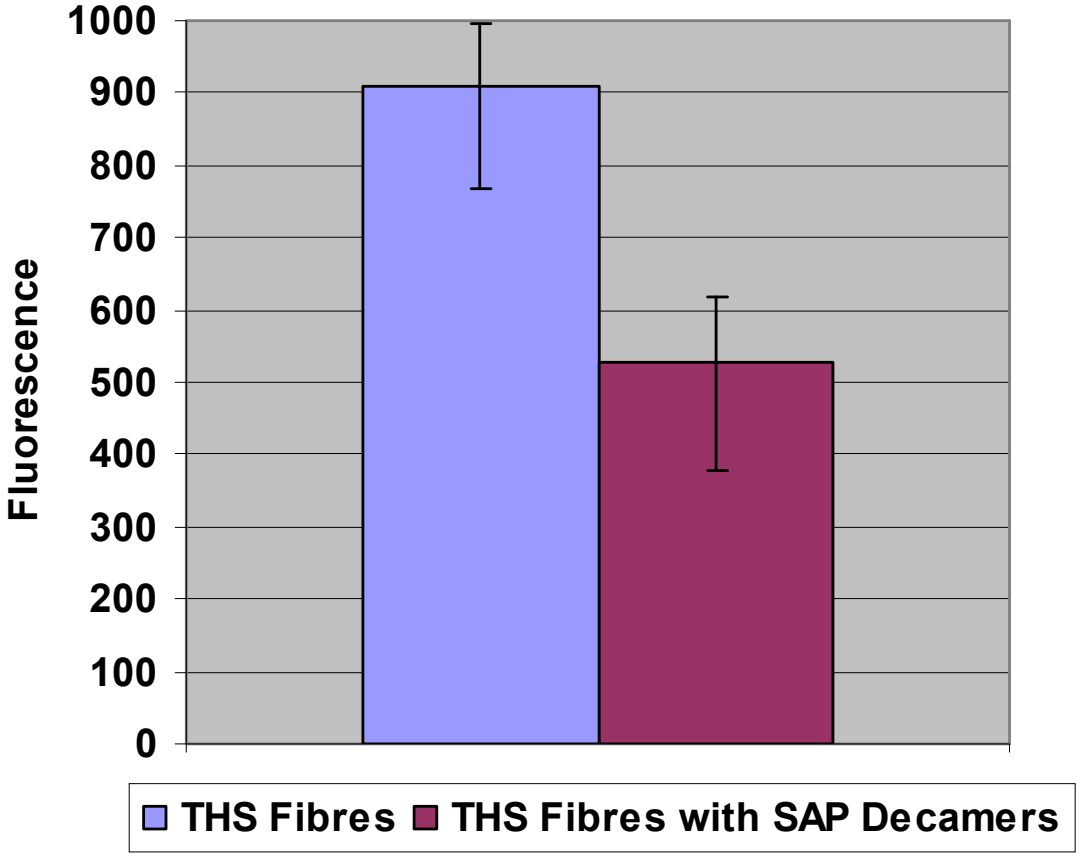


Figure 31: A bar graph comparing average ThT fluorescence of insulin fibres grown in the presence/absence of 0.2mg/ml SAP pentamers in high salt (THS) buffer. Fibres were rotated for 10 days on a revolving Stuart Rotator SB3 at 20rpm at room temperature. Each bar is an average of three samples and has error margins included to show the range of those three samples. The fibres grown in THS buffer produce an average fluorescence of 574.3 where as those grown in the presence of SAP decamers produce an average fluorescence of 347.3. The presence of SAP decamers causes an average decrease in fibre formation of 39.5%.

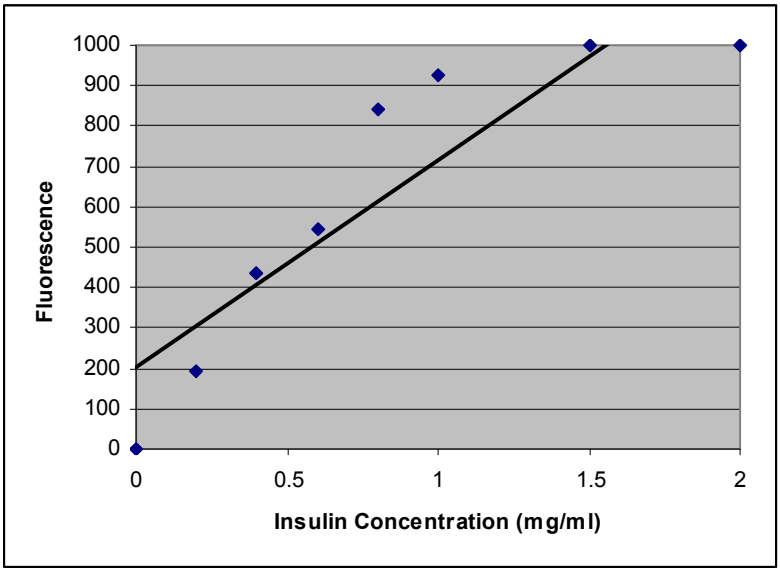


Figure 32: Graph showing the correlation between insulin concentration and ThT fluorescence. A strong positive correlation of 0.897 is observed. This illustrates that the more protein present, the higher amount of cross β-sheet fibres produced and thus the higher the ThT fluorescence.

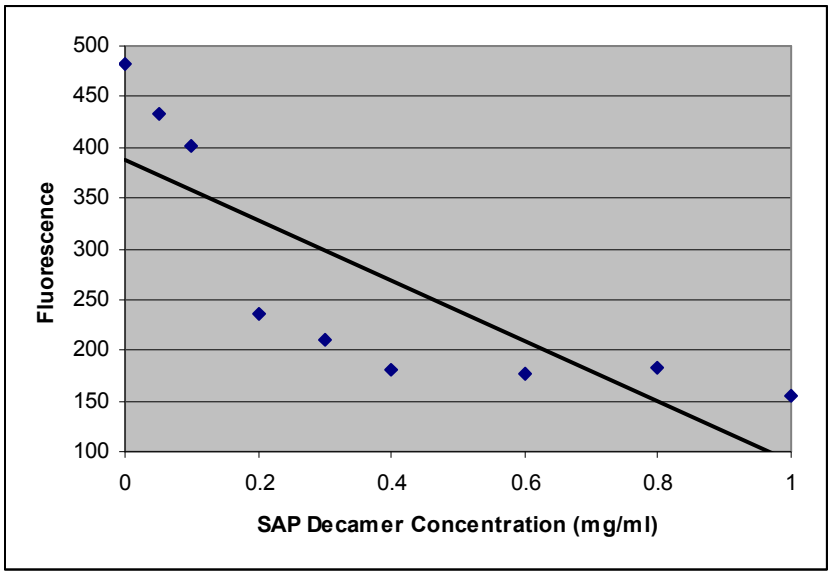


Figure 33: Graph showing the correlation between SAP decamer concentration and ThT fluorescence. A strong negative correlation of -0.812 is observed. This illustrates that the more SAP decamers present, the more fibre amounts are reduced thus the lower the ThT fluorescence.

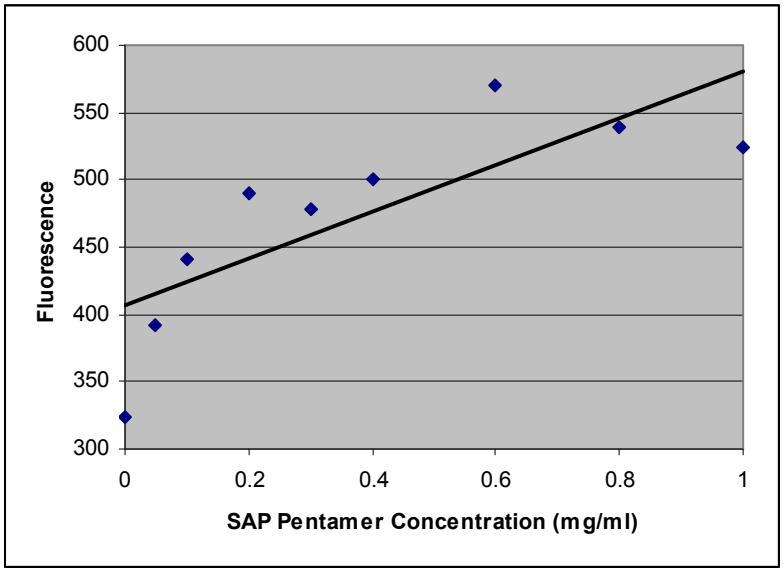


Figure 34: Graph showing the correlation between SAP pentamer concentration and ThT fluorescence. A strong positive correlation of 0.713 is observed. This illustrates that the more SAP pentamers present, the more fibre amounts are increased thus the higher the ThT fluorescence.

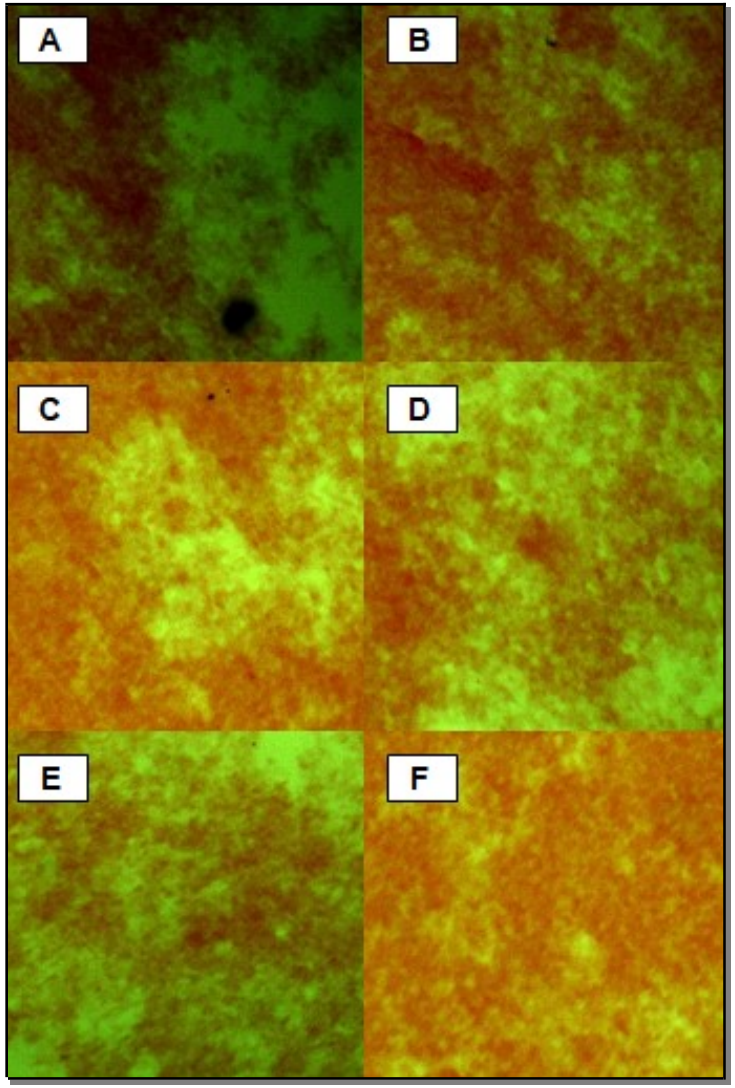


Figure 35: Images of insulin fibres stained with Congo red. All samples showed the apple green birefringence typical of cross-β sheets under polarised light. The images were viewed using a Zenith Microlab ML-1000M Monocular Laboratory Microscope 60040 at a magnification of x40for samples A, B, E, F and x100 for samples C,D. Sample contents are as follows: A – Insulin fibres grown in TE buffer. B) – Insulin fibres grown in TE buffer with SAP decamers. C – Insulin fibres grown in TC buffer. D) – Insulin fibres grown in TC buffer with SAP pentamers. E) – Insulin fibres grown in THS buffer. F) – Insulin fibres grown in THS buffer with SAP decamers.

In the samples grown in TE buffer – A and B – a decrease in fibre density is observed in B, the sample containing SAP decamers. In the samples grown in TC buffer – C and D – an increase in fibre density is observed in D, the sample containing SAP pentamers. And in the samples grown in THS buffer –E and F – a decrease in fibre density is observed in F, the sample containing SAP decamers There are no significant differences seen between the images of the TE and THS samples indicating minimal impact of the salt on the fibres.

This is a reassuring result. It can now be presumed that all differences with regard to fluorescence noted within the confines of the TC buffer assays are due to the change in SAP conformation alone and not the involvement of salt. The additional calcium variable does not appear to affect the relationship between SAP pentamers and fibres either; however it does appear to be responsible for reduction of total fibre amounts produced. This means that direct comparisons between the TC and TE assays cannot necessarily be made, however, proportional figures and ratios can be discussed and analysed.

Another observation became apparent over time. The control fibres - those grown in the high salt conditions of the THS buffer - were less stable than the original TE fibres. ThT fluorescence of the THS fibres would deplete to 25% of their original amount within 2 to 3 months whereas TE fibres would remain stable for up to a year before a similar depletion was observed. It would appear that the salt has no observable effect on fibre growth but instead is responsible for a lack of stability on those resulting fibres.

That being stated, THS fibres which were grown in the presence of SAP 'decamers' drastically increased in stability, withstanding up to 8 or 9 months before depletion. A similar pattern was experienced with the TE fibres. It seems that SAP is able to add stability to fibres even when in its decameric form and acting as a chaperone. Perhaps in the decameric form enough conformational changes have occurred in order to allow fibre:SAP binding of a different kind. Perhaps the edge of the decamer is interacting with the fibre instead. The increased stability of insulin fibres grown in TE buffer with the presence of SAP decamers compared to those grown in the same conditions in the absence of SAP decamers would certainly suggest so.

That being said, it is important to remember that SAP, although predominantly decameric in this TE buffer, may also be available in its pentameric state as SAP never exclusively exists as decamers according to Thompson *et al.* in 2002. It is a mixture. The most logical conclusion, therefore, would be that when in its "decameric" state, it is the predominant decameric molecules

which are acting as chaperones, refolding available monomers and the subsidiary pentameric molecules are binding to the fibres present in solution and increasing their stability.

SAP Control

When submit to the same inversion conditions as the prepared insulin, lone SAP - at concentration 0.2mg/ml - did not form fibres or any other form of amorphous aggregate; it remained soluble in solution. This demonstrates that SAP concentration and function does not diminish during the course of the rotator assay and it should not interfere with ThT fluorescence or any other analytical methods for cross β-sheets.

The Role of SAP in Fibrillogenesis

Thus far from this chapter the following conclusions can be made: In the case of insulin, SAP appears to have a dual role in fibrillogenesis; it can enhance or reduce fibre formation depending on its conformation. When in its decameric state SAP inhibits fibre formation and decreases total fibre mass. When in its pentameric state SAP acts as a fibrillogenic stabiliser or enhancer. This dual behaviour of SAP is consistent and is not affected by environmental factors such as rotator speeds, concentration changes or high salt concentrations. SAP alone, when submitted to the same conditions as the prepared insulin, does not form amorphous aggregate nor cross-β sheets; it remains in its solubilised state as confirmed by ThT fluorescence. This confirms that all cross β-sheet fibres are purely insulin in origin.

SAP State	Available face	Role in fibrillogenesis
Decamer	A-face	Molecular chaperone
Pentamer	B-face	Fibrillogenic catalyst/enhancer

A table to illustrate the roles of each SAP conformer.

The existing literature has already demonstrated increased fibre stability through SAP binding, and so provides one explanation as to why fibre amounts increase in the presence of SAP pentamers. The increased stability would encourage the rope-like structures to extend and therefore increase the amount of fibre present in the solution. The behaviour of the decamers, however, is less clear. Is the diminishing function carried out through partial binding of unfolded species or through the depletion of existing fibres? The former would seem more feasible due to a number of reasons: Firstly, during growth, a high number of partially unfolded intermediates are available for decamer interaction. Secondly, cross β-sheet fibres are extremely stable so detachment of monomers from the core via the decamers seems wholly unlikely. Finally, when in its decameric form, the binding (B) side of SAP is involved in an interaction with another SAP molecule so is not actually available for fibre interaction via its binding (B) face. That being said, alternate interactions via the edge of the decamer cannot be ruled out.

When comparing SAP decamers to existing chaperones quite a few similarities are observed. The decamer exists as a large cylindrical structure formed of two rings of oligomers. This description is similar to that of chaperonins. Although not as big as the GroEL chaperonin, for example, 250kDa rather than 798kDa, the SAP decamer also has a hollow chamber which passes through its centre. The larger GroEL structure often deals with proteins up to 60kDa in size. Being much smaller than GroEL, it is doubtful that SAP would be able to interact with proteins up to this size. However, in this case, it is the insulin protein which is involved. Insulin is only 6kDa so could quite easily enter one of the SAP rings, bind to the inner wall via

hydrophobic interactions providing enough conformational changes to encourage protein folding. Not only is it entirely possible that SAP is indeed some form of chaperonin but it is able to act as so even without the need of ATP.

Fibre Comparison using Congo Red

As with the ThT results, the effect of SAP on fibres can be distinguished using Congo red images. The following samples underwent Congo red analysis and were compared in Figure 35:

- Insulin fibres grown in TE buffer
- Insulin fibres grown in TE buffer with the addition of 0.2mg/ml SAP decamers
- Insulin grown in TC buffer
- Insulin grown in TC buffer with the addition of 0.2mg/ml SAP pentamers
- Insulin grown in THS buffer
- Insulin grown in THS buffer with the addition of 0.2mg/ml SAP decamers

All samples showed the apple green birefringence typical of cross-β sheets under polarised light. The images were viewed using a Zenith Microlab ML-1000M Monocular Laboratory Microscope 60040 at a magnification of x40 except for those grown in TC buffer with and without SAP. This was due to the fact that these particular samples had less clustering of apple green birefringence. A higher magnification of x100 was required for comparison to the previous samples.

Image A and B represent insulin fibres grown in TE buffer without and with SAP decamers respectively. As expected, image B – those grown in the presence of SAP decamers – demonstrate a decrease in fibre density

compared to those grown without. This further confirms the chaperoning activity of the decamer molecule. The decrease in fibres cannot be given a value in this case as the Congo red technique is not a quantitative method. Images C and D represent fibres grown in TC buffer without and with SAP pentamers respectively. Here, an increase in fibre density is witnessed in image D – those grown in the presence of the SAP pentamers thus confirming their activity as an enhancer and stabiliser. Finally, images E and F represent insulin fibres grown in THS fibres without and with the presence of SAP decamers respectively. Again, a decrease in fibre density is seen in F; those grown in the presence of SAP decamers. This result confirms that the chaperoning activity of the SAP decamers is not altered by the high salt conditions.

No significant differences are seen between the images of the TE and THS samples. This clarifies that the high salt concentrations do not have a substantial impact on the amount or density of fibres made. These results, alongside the ThT results, provide a strong indication that all samples had indeed formed fibres and not amorphous aggregate. However, neither the ThT or Congo red techniques provide enough information regarding the structure of the samples in order to confirm this. Also, having grown in varying conditions it seemed unlikely that the morphologies of each sample would be completely identical even though they originated from the same stock solution of insulin. Structural work was required.

Fibre Comparison using Transmission Electron Microscopy (TEM)

Initial structural analysis was carried out using transmission electron microscopy (TEM). Although the specific cross β-sheets fold cannot be seen using TEM due to magnification limitations, the typical rope like structure *can* be distinguished. The visualisation of rope like structures in TEM images not

only confirms the presence of the cross β-sheet structure specific to amyloid but also provides a comparable image with regards to structure.

The following samples were analysed using TEM and compared **()**:

- Insulin fibres grown in TE buffer
- Insulin fibres grown in TE buffer with the addition of 0.2mg/ml SAP decamers
- Insulin grown in TC buffer
- Insulin grown in TC buffer with the addition of 0.2mg/ml SAP pentamers
- Insulin grown in THS buffer
- Insulin grown in THS buffer with the addition of 0.2mg/ml SAP decamers

The TEM images of fibres grown in the TE buffer exhibited morphology typical of cross-β sheet fibres. At 140,000x magnification, rope-like structures approximately 7nm in diameter and of indefinite length are clearly visible. A similar result is also seen in the TE fibres grown in the presence of SAP decamers. Very little difference is seen between the two. This would indicate that the decamers, although involved in growth rates and total fibre amounts, have no (or very little) influence on fibre structure and morphology. This would also indicate that the decamers are not physically bound to the fibres when in their decameric state. SAP decamers have a depth of 70Å or 7nm and a diameter of 100 Å or 10nm. If it were true that the decamers were bound to the fibres, there would be a large increase in diameter of the fibre structures. This further suggests the theory that the SAP interacts with the fibre in a mechanism similar to a chaperonin molecule.

Contrastingly, the fibres grown in the calcium containing (TC) buffer are rather different. Those grown in the absence of SAP have a slate-like morphology. They are much wider in diameter than previously seen - 30nm rather than the expected 10nm – and also much shorter growing to only 100nm in length. The

presence of salt or calcium ions in the buffer has undoubtedly had a marked effect on fibre structure and would explain the instability noted on these fibres. Remarkably, the fibres grown in the TC buffer in the presence of 0.2mg/ml SAP pentamers have developed a structure much like the typical rope-like morphology with a wider diameter of 15-20nm and a length of up to 500nm. These results suggest that the SAP pentamers are playing an active role in fibre assembly. Upon binding, they appear to direct and encourage the formation of a more stable formation however they are unable to completely counteract the effect that either the salt or the calcium is having on the structure.

These increases in diameter could be due to one of two reasons. The first being the SAP pentamers are bound to the fibres. Although completely plausible, stacks of SAP cannot be distinguished in either of these images so is likely not the case. Alternatively one could suggest that the increased diameter in both the TC images is a result of the insulin monomers interacting differently with one another in order to create the fibres. Cross β-sheet formation occurs when, in this case, the insulin molecule partially unfolds and interacts with another of its kind. Perhaps when subject to different environmental conditions, the monomers unfold in a different place or to a different degree. In turn, this would cause the monomers to interact differently upon cross β-sheet formation and cause smaller or larger side bulks depending on the degree of unfolding. The more unstable the environment, the larger the degree of unfolding and thus the wider the fibre will be when stacked. The more stable the environment, the smaller the degree of unfolding and thus the thinner the fibre will be.

The control fibres, *i.e.* the fibres made in THS buffer closely resemble the typical rope-like structures as seen in the TE buffer. The fibres are approximately 7nm in diameter and are of indefinite length. This establishes that it is the calcium which causes the dramatic change of morphology to the ‘slate’ pattern and not the high salt conditions. However, a distinct difference is noted between the TE and THS fibres. The THS fibres appear slightly fragmented and have a ‘sickly’ quality to them. This suggests that the salt

does have a slight impact on fibre structure which explains why these fibres are less stable than the TE fibres. THS fibres grown in the presence of 0.2mg/ml SAP lack this ‘sickly’ quality, further demonstrating SAPs stabilising qualities. SAPs fibrillogenic involvement is clearly structural as well as mechanical.

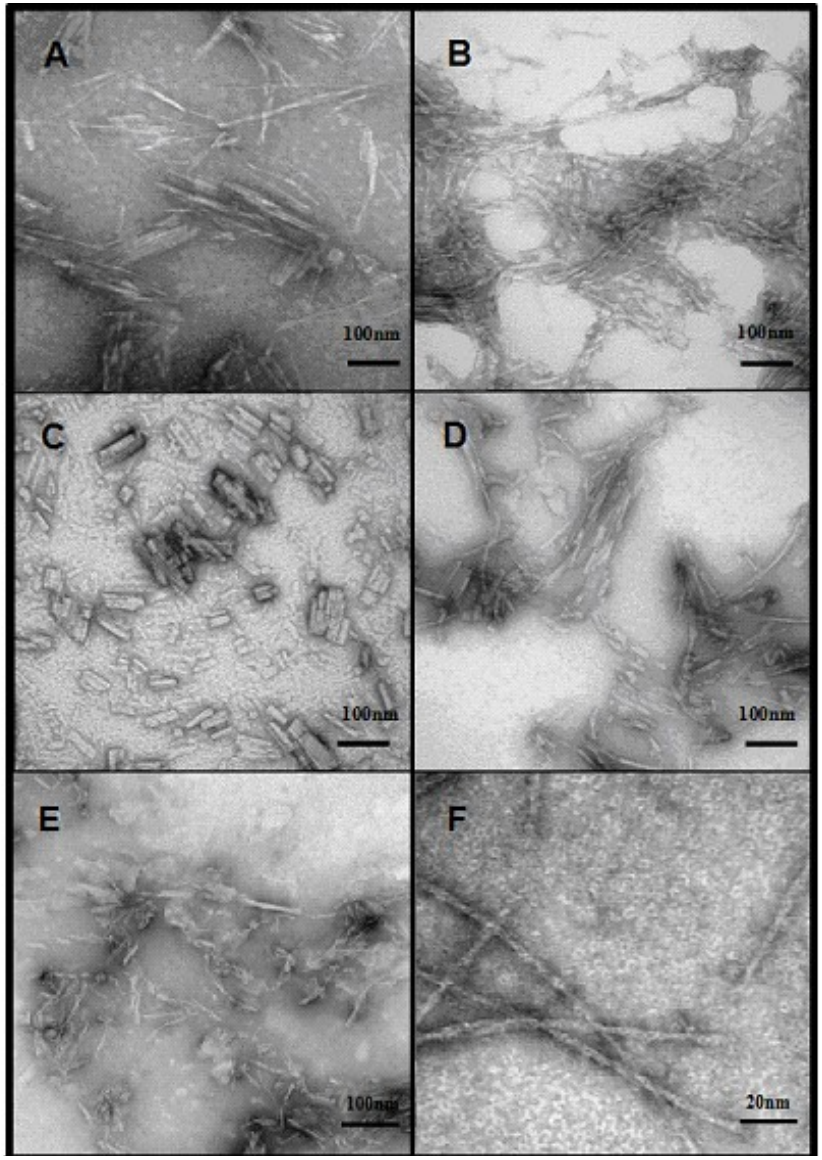


Figure 36: TEM images of insulin fibres. Images were obtained using a Philips (FEI) CM120 TEM microscope at a magnification of 140,000x. A) TE fibres grown in the absence of SAP. Typical rope-like structures approximately 7nm in diameter and of indefinite length are seen. B) TE fibres grown in the presence of SAP. Here, the rope-like structures are approximately 8.5nm in diameter and of indefinite length. C) TC fibres grown in the absence of SAP have a different morphology than expected. They have a slate-like morphology 30nm in diameter and approximately 100nm in length. D) The TC fibres grown in the presence of SAP pentamers were again the typical rope-like morphology but were slightly wider and shorter than expected. 15-20nm in diameter and approximately 500nm in length. E) The high salt conditions created 'Sickly' THS fibres in the absence of SAP. F) 700,000x magnification of THS fibres grown in the presence of SAP decamers show fibres have reverted back to the typical rope-like structures approximately 7nm in diameter. SAP pentamers can be seen in the background and attached to the fibres in this image.

Due to the higher magnification of the THS fibres grown with SAP images, the SAP structure can also be seen. The distinctive disc shaped structure is scattered in the background of the image but also appears to be stacked on the amyloid fibres. This illustrates that SAP does indeed have interactions with the fibres themselves and not just with the amyloid prone intermediates. It should be noted that the staked SAP structures appear to be in the pentameric form. To explain this one must remember that SAP never exists solely as a decamer but instead as a mixture of decamers and pentamers. This lack of evidence regarding decamer binding to fibres would suggest that it is indeed the pentamers which bind to the fibres and in doing so increase stability and enhance growth, however, the decamers carry out their chaperoning activity in solution, separate from the fibres.

Considering there is a far greater number of non-bound decamers than bound pentamers it would explain why when added as a 'decamer' *i.e.* a mixed solution, the decameric activity is the dominant form. Although the pentamer is present, its low numbers are unable to maintain or exceed the rate of reaction of the decamer activity. Whether there is a condition which allows pure decamers to form is yet to be seen but this seems unlikely for two reasons. The first being SAPs tendency to exist exclusively as decamers in the serum under conditions which would normally promote decamer formation. Clearly this is the favoured formation whether due to stability or binding ability. Secondly, due to the apparent reversible nature of the decamer interactions it would be unlikely that reasonable conditions could be found to encourage full saturation of decamer binding. even if such conditions could be found, they would not hold any sense of resemblance to in vivo conditions and would never be able to be replicated in the body so is irrelevant to these particular experiments.

Fibre Comparison using Circular Dichroism

The following samples were processed using CD and compared:

- Insulin fibres grown in TE buffer

- Insulin fibres grown in TE buffer with the addition of 0.2mg/ml SAP decamers
- Insulin grown in TC buffer
- Insulin grown in TC buffer with the addition of 0.2mg/ml SAP pentamers
- Insulin grown in THS buffer
- Insulin grown in THS buffer with the addition of 0.2mg/ml SAP decamers
- Native insulin (before and after dialysis)
- Native SAP
- Aggregated SAP

In the cases where SAP was involved, SAP was removed from the solution using EGTA and washed away. Although EGTA was not necessary for the removal of decamers from solution as they, unlike the pentamers, do not experience metal specific binding to the fibres, it was decided that all samples should be submitted to the same preparatory conditions in order to ensure a fair test. Each sample was transferred to water, a non chiral solution which would not affect the CD results, and diluted to the same optical density. It was chosen to dilute the fibres to the same optical density rather than the same concentration due to the impracticalities of measuring the concentration of a solid. CD measurements were obtained at Diamond Light Source, Oxfordshire, on beamline 23 (B23). Each sample was measured a total of three times in order to obtain an average.

As expected, the native insulin samples demonstrate curves typical of alpha helices **(Figure 37)**. The typical peak and double depression are seen in both spectra with little difference seen between the dialysed and non-dialysed samples. The x axis interception point occurs at wavelength 203nm for both where as the wave amplitudes differ slightly at +34 and -16mdeg for the non-dialysed sample and +32 and -14mdeg for the dialysed sample. As explained in the introduction, insulin is a predominantly helical structure composed of two chains. The shorter A chain contains an N-terminal helix linked to an anti-parallel C-terminal helix and the longer B chain has a central helical segment. This explains why a helical spectrum is produced. In contrast to this, SAP is formed predominantly of ß-sheets, ß-turns and some unknown structures. The

CD data of native SAP illustrates this by producing a curve typical of β-strands. The single peak and depression typical of ß-structures is observed with the shift of the x axis interception point to the right compared to that of the helices. The x axis interception point for native SAP occurs at 207nm with amplitude peaking at +26 and -13mdeg. When made into amorphous aggregate through the addition of a few drops of 37% hydrochloric acid, the spectra changes. A very low and broad peak and depression close to the base line is produced. It has a shifted x axis interception point of 204nm. This shows that some level of β-strand has been retained in SAP in acidic conditions but the majority of secondary structure has been lost.

Upon submitting the different fibre samples to CD, various observations were noted **(Figure 38, Figure 39)**. First and foremost, each fibre spectra produced a curve typical of β-strands. This shows that the rotator assays had caused the insulin samples to undergo a huge conformational change from predominantly α-helices to predominantly β-sheets. However, each of the produced curves intercepted the x axis at a different point. These unique interception points imply that each fibre is cross β-sheet in nature but slightly unique in morphology. This is in agreement with the TEM results. The CD spectra of the fibre samples have all shifted slightly towards the right hand side of the graph compared to native insulin. Interception points of the x axis are no longer at 203nm but rather at unique points of 217nm (TE fibres), 210nm (TE fibres + SAP decamers), 221nm (TC fibres), 213nm (TC fibres + SAP pentamers) 208 (THS fibres) and 211 (THS fibres +. SAP decamers). Had all fibres contained the same ratio of secondary structure *i.e.* if their morphology were identical, the spectra would intersect the x axis at the same point regardless of concentration as per the two native insulin spectra.

The unusual result from the CD data is that each fibre curve is of a different amplitude compared to its' peers. There is a range of +2 to +25mdeg in the positive peak and -2 to -16mdeg in the negative depression with the TE fibres having the highest amplitude and the TC fibres having the lowest. The amplitude of a curve or spectra is representative of the protein concentration of the sample submitted to CD analysis. Taking into consideration that each

sample was diluted to the same optical density, it is logical to expect each sample to have the same concentration of protein thus the same amount of cross β-sheets and equal amplitude. This is clearly not the case.

The amplitude differences of each curve indicate that the amount, or mass, of cross β-sheets in each sample is different. Yet, each sample had the same optical density when undergoing CD analysis. This can only be explained one way. There is more than one component present in each sample. This extra component or components are having an effect on fibre concentration. It is safe to say that native insulin is not responsible for this impact on the results as it was completely removed during sample preparation. Even if it were present, it is logical that each sample would have a proportional amount of native insulin thus making any affect it had on the spectra a uniform effect.

It is possible that SAP is having an effect on the results. Each fibre, CD sample subsequently underwent SDS-PAGE electrophoresis. The results show a faint band at a molecular weight of approximately 25kDa in each of the samples which originally contained SAP; the equivalent of a single SAP domain. Clearly the EGTA was unable to fully detach SAP form the fibres. It is indeed possible that the SAP could have had some influence on the CD spectra causing anomalous results to arise. However, regardless of this being true of not, this would not explain the difference in magnitudes between the samples which didn't contain any SAP to begin with. It cannot be the SAP which is having this drastic effect on amplitude. Something else must be causing this phenomenon. Another factor is causing the optical densities to remain constant while the β-sheet concentrations vary **(Figure 40)**.

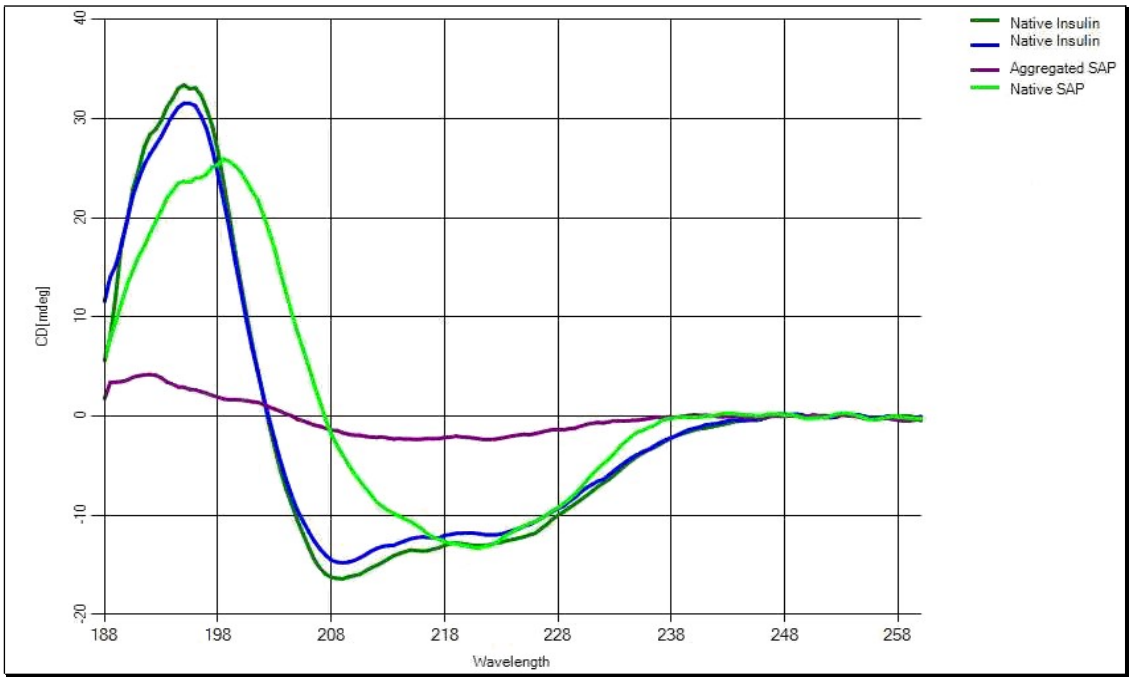


Figure 37: Control CD spectrum showing insulin and SAP in their native forms. Measurements were taken at Diamond Light Source, Oxfordshire, on beamline 23 (B23) – soft condensed matter. Both native insulin samples – before (dark green) and after (blue) dialysis - produced spectra typical of α-helices crossing the x-axis uniformly at 202nm. Native SAP (bright green) produced spectra typical of β-sheets with an x-intercept of 207nm. Aggregated SAP (purple) produced spectra of no defined structure.

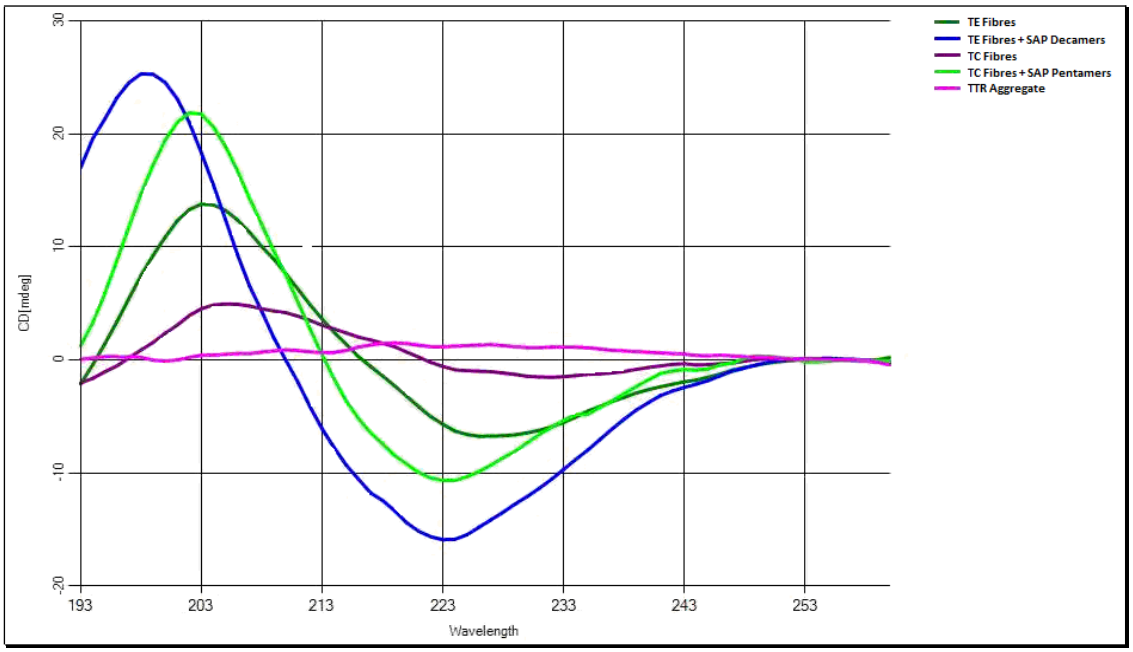


Figure 38:Average CD results of TE and TC insulin fibres grown in the presence and absence of SAP. TE fibres are represented in dark green, TE fibres + SAP are represented in blue, TC fibres are represented in dark purple, TE fibres + SAP are represented in bright green and TTR aggregate is represented in pink. Each spectra is an average of ten measurements and was taken at Diamond Light Source, Oxfordshire, on beamline 23 (B23) – soft condensed matter. All insulin fibres produced spectra typical of β-sheets. The differences in amplification are due simply to differences in fibre concentrations. However, the differences in x-axis interception would indicate that the structure of each fibre is different. X-axis interception ranges from 210-222 nm. Those fibres grown in the presence of SAP have higher amplification and x- axis interception shifted to the left of the graph.

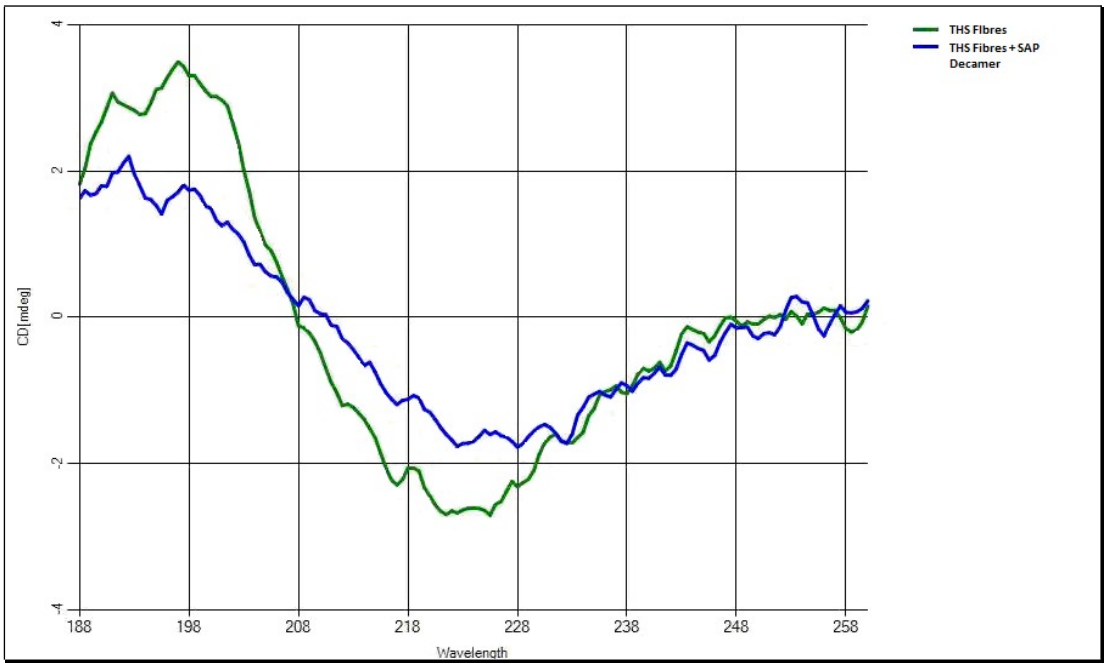


Figure 39: Average CD results of THS fibres with/without SAP decamers. THS fibres are represented in green and THS fibres + SAP decamers are in blue. Each spectra is an average of ten measurements and was taken at Diamond Light Source, Oxfordshire, on beamline 23 (B23) – soft condensed matter. Although on a small scale, the curves are still typical of β-sheets with interception at 208 and 211nm.

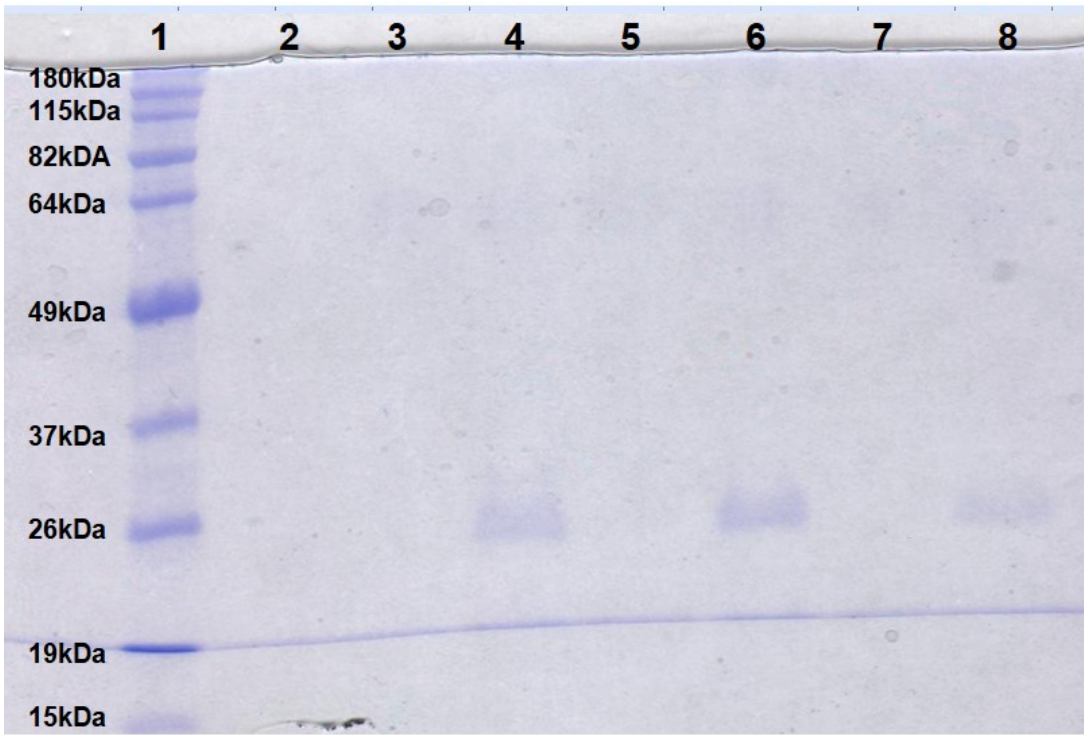


Figure 40: Electrophoresis gel showing of the six samples submitted to CD analysis following EGTA washes and transference to water. The contents of each lane are as follows: 1) Pre-stained molecular marker – weights are labelled 2) Blank 3) TE fibres 4) TE fibres + SAP decamers 5) TC fibres 6) TC fibres + SAP pentamers 7) THS fibres 8) THS fibres + SAP decamers. The presence of faint bands at 25kDa in lanes 4, 6 and 8 indicate that SAP was still attached to fibres even after washing with EGTA.

Optical density is the measurement of how much a refractive medium retards transmitted rays of light. It takes into account the mass or density of the medium but not the composition. The only logical explanation for this strange occurrence is that all the fibre samples did indeed contain the same amount of precipitate within as measured by the optical density but that the precipitate was not purely cross β-sheets. This suggests that each precipitate consisted of at least two insoluble substances. One of these is cross β-sheets and the second must be amorphous aggregate.

As shown in Figure 37 with the aggregated SAP, amorphous aggregate does not produce a peak on the CD spectra as it lacks secondary structure. The presence of amorphous aggregate in a cross β-sheet fibre sample shouldn't alter results in a drastic way although it is possible that it could cause the x axis interception to shift slightly. However the presence of amorphous aggregate would defiantly cause discrepancies to occur between concentration of cross β-sheet fibres and the optical density of the overall sample. Each sample may have the same amount of aggregate mass but a different cross β-sheet: amorphous aggregate ratio. It is not possible to separate the two as both are white solids with similar densities and made of the same protein. However, seeing as the presence of the amorphous aggregate is not affecting the CD results this is not something to be concerned about. That being said, this observation presents a new set of questions. Why is amorphous aggregate formed? How much is formed? Does this vary between the fibres grown in different conditions?

Further scrutiny of all the CD results was achieved using the programme 'CD Apps'. This programme – which was developed at Diamond Light source, beamline B23 – takes the spectra produced from CD analysis and converts the resulting curves into unique numerical values. These values represent the ratio of α-helices and β-strands present in a particular sample as a percentage. By using this programme one can see the secondary structure composition in ore detail.

Each of the fibre sample spectra was submitted to this programme alongside native insulin as a comparison and the results are represented in the table below:

Secondary Structure	Native Insulin	TE Fibres	TE Fibres + SAP Decamers	TC Fibres	TC Fibres + SAP Pentamers	THS Fibres	THS Fibres + SAP Decamers
Alpha Helix	15.6%	0.3%	19.7%	0%	0.1%	0.1%	0.1%
Distorted Alpha Helix	14.3%	6.3%	15.9%	4.5%	4.0%	3.7%	4.0%
Alpha Helix (both types)	**29.9%**	**6.6%**	**35.6%**	**4.5%**	**4.1%**	**3.8%**	**4.1%**
Beta Strand	15.6%	25.3%	22.4%	26.7%	30.1%	28.6%	28.7%
Distorted Beta Strand	8.2%	14.6%	11.5%	13.7%	14.4%	14.0%	13.9%
Beta Strand (both types)	**23.8%**	**39.9%**	**33.9%**	**40.4%**	**44.5%**	**42.6%**	**42.6%**
Unordered	**46.3%**	**53.5%**	**30.5%**	**55.1%**	**51.4%**	**53.6%**	**53.3%**

Table 5: A table comparing the secondary structures of each fibre sample calculated using the 'CD Apps' programme available at Diamond Light Source, beamline B23.

When looking at the native sample you can see that insulin is normally constructed of 29.9% α-helices, of this, almost half (14.3%) are distorted. A further 23.8% of the structure is made of β-strands (8.2% distorted) whist the remaining 46.3% represents turns and unordered structures. To put these numbers into perspective, SAP, a predominantly β- stranded structure, consists of 21.8% α-helices (6% distorted), 30.1% β- strands (11.4% distorted) and the remaining 48.1% is turns and unordered structures. Alternatively the aggregated SAP– achieved by exposing SAP to hydrochloric acid solution – contains 10% α-helices with 90% dedicated to unordered structures.

From these results it would appear that a protein must consist of approximately 30% of either helices or strands in order to be considered predominantly of that form. Insulin is constructed of 29.9% α-helices and SAP is 30.1% β- strands. Whether these secondary structures are distorted or not appear to have no effect on the secondary structure itself but most likely will alter the packing of the tertiary structure.

In a typical aggregated structure, *i.e.* in the amorphous state, a protein is able to maintain some form of secondary structure but the majority is unordered such as with the aggregated SAP mentioned above. This is not the case, however, when dealing with cross β-sheet aggregation. Cross β-sheet structures do have a secondary structure. Each of the fibres samples under the scrutiny of the CD Apps programme have decreased in α-helix percentage with the exception of the TE fibres + SAP decamers. The drop from 29.9% to 6.6% (TE fibres), 4.5% (TC fibres) 4.1% (TC + SAP pentamers) 3.8% (THS fibres) and 4.1% (THS fibres + SAP decamers) in each of the samples clearly shows that the primary secondary structure is lost. The variety on the degree of loss can be explained by the variation of the conditions to with the samples were submit to originally. Oddly, the percentage of α-helices in TE fibres grown in the addition of SAP decamers has increased to 35.7% however, as with the other samples, the percentage of β-strands has also increased. The percentage has doubled to 39.9% (TE fibres), 33.8% (TE fibres + SAP decamers), 40.4% (TC fibres) 44.5% (TC fibres + SAP pentamers) 42.6%

(THS fibres) and 42.6% (THS fibres +SAP decamers) with up to ¾ of these being ordered. Although the TE fibres were able to maintain (and even increase) their helical structure, the β-sheet conformation has a higher percentage so is dominant. If anything, this result shows how little the insulin needs to unfold before this drastic change in conformation occurs in optimum conditions.

Fibre Comparison using Bicinchoninic Acid Assay (BCA)

Fibres grown in the following six conditions were processed using the BCA assay and compared. Four different samples per condition were analysed in order to obtain an average:

- Insulin fibres grown in TE buffer
- Insulin fibres grown in TE buffer with the addition of 0.2mg/ml SAP decamers
- Insulin grown in TC buffer
- Insulin grown in TC buffer with the addition of 0.2mg/ml SAP pentamers
- Insulin grown in THS buffer
- Insulin grown in THS buffer with the addition of 0.2mg/ml SAP decamers

BCA is a method used to determine the protein concentration of a solution between the concentrations 0.5 µg/ml to 1.5 mg/ml. It incorporates all protein present including solubilised protein and aggregated protein whether present as fibres or amorphous aggregate. By removing the solubilised protein from each sample, a comparison between the BCA and ThT results could be made and an indication of the fibre:amorphous aggregate proportions could be determined.

TE fibres grown in the absence of SAP had a direct positive correlation between protein precipitation and ThT signal – *i.e.* the more protein used the

more fibre that was made. Of which, approximately 80% of the insulin was made into cross β-sheets, 16% was amorphous aggregate and the remaining 4% continued as native insulin or soluble intermediates. However, the TE fibres grown in the presence of SAP decamers did not follow this pattern. They had a lower ThT signal yet contained almost equal amounts of precipitate as their SAP free counterparts. Only 20% of the insulin became cross β-sheet whereas 62% became amorphous aggregate with the remaining 18% continuing as either native insulin or soluble intermediates. This suggests that the SAP decamers promote amorphous aggregate formation. Alternatively it would be sensible to suggest that the decamers promote refolding of unfolded aggregates **(Figure 41)**.

The samples grown in TC buffer, like their TE equivalents, also experienced direct correlation between protein precipitation and ThT signal; the more protein, the more fibre that was made. Approximately 21% of the insulin was made into cross β-sheets, 69% was amorphous aggregate and the remaining 10% continued as native insulin or soluble intermediates. These lower amounts are most likely due to the presence of calcium and salt in the buffer creating a less stable fibre structure. Again, those grown in the presence of SAP - in this case SAP pentamers – produced different results. Those samples grown in the presence of SAP pentamers exhibited higher ThT signals yet contained less overall precipitate compared to their control equivalents. Approximately 41% of the insulin was made into cross β-sheets, 57% was amorphous aggregate and the remaining 2% continued as native insulin or soluble intermediates. This supports the theory that SAP pentamers are promoting the formation of amyloid fibres.

Perhaps SAP is the regulator of a reversible pathway between cross β-sheet fibres and amorphous aggregate **(Figure 42)**. The decamers would therefore be responsible for shifting equilibrium towards amorphous aggregate whereas the pentamers are responsible for shifting equilibrium towards cross β-sheets. In fact, it is even more plausible to suggest that the role of SAP is to maintain equilibrium between cross β-sheet fibres and the native structures as the body has no use for amorphous aggregate (further discussions below).

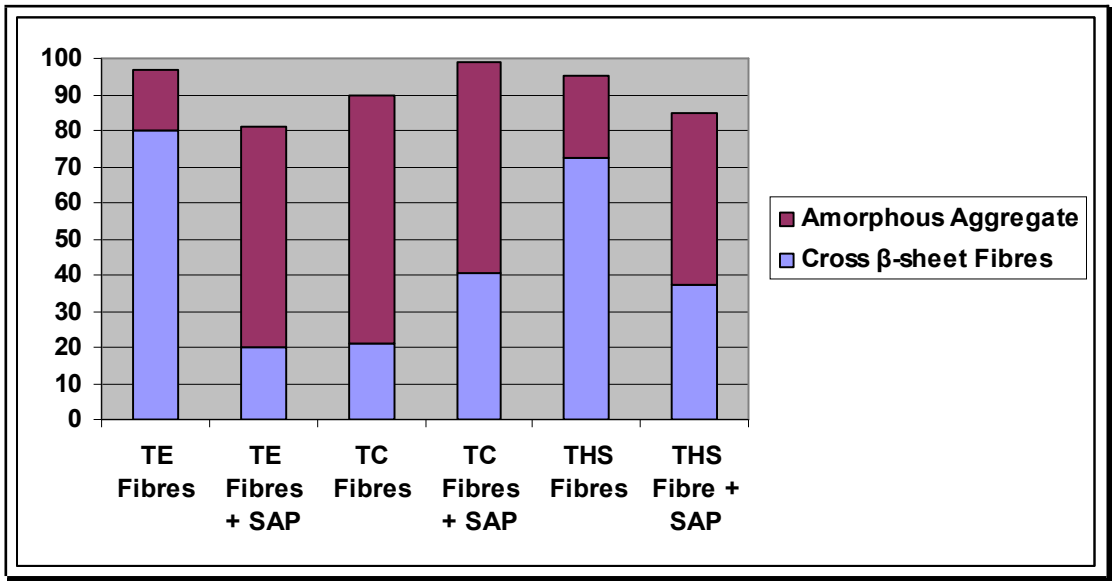


Figure 41: A bar graph comparing composition of the insulin precipitates made from different rotator conditions. Amorphous aggregate is represented in purple and cross β-sheet fibres are represented in blue. The precipitates were processed using the BCA assay and compared to their ThT fluorescence signal in order to obtain the following figures. Four different samples per condition were analysed in order to obtain an average. In TE fibres approximately 80% of the insulin was made into cross β-sheets, 16% amorphous aggregate and the remaining 4% continued as native insulin or soluble intermediates. TE fibres grown in the presence of SAP decamers had 20% become cross β-sheet, 62% amorphous aggregate and 18% continuing as either native insulin or soluble intermediates. TC buffer fibres produced 21% cross β-sheets, 69% amorphous aggregate and 10% as native insulin or soluble intermediates. TC fibres grown in the presence of SAP pentamers produced 41% cross β-sheets, 57% amorphous aggregate and 2% continued as native insulin or soluble intermediates. THS fibres produced 73% cross β-sheets, 22%, amorphous aggregate and 5% continued as native insulin or soluble intermediates. THS fibres grown in the presence of SAP decamers yielded 37% cross β-sheets, 48% amorphous aggregate with 15% continuing as either native insulin or soluble intermediates.

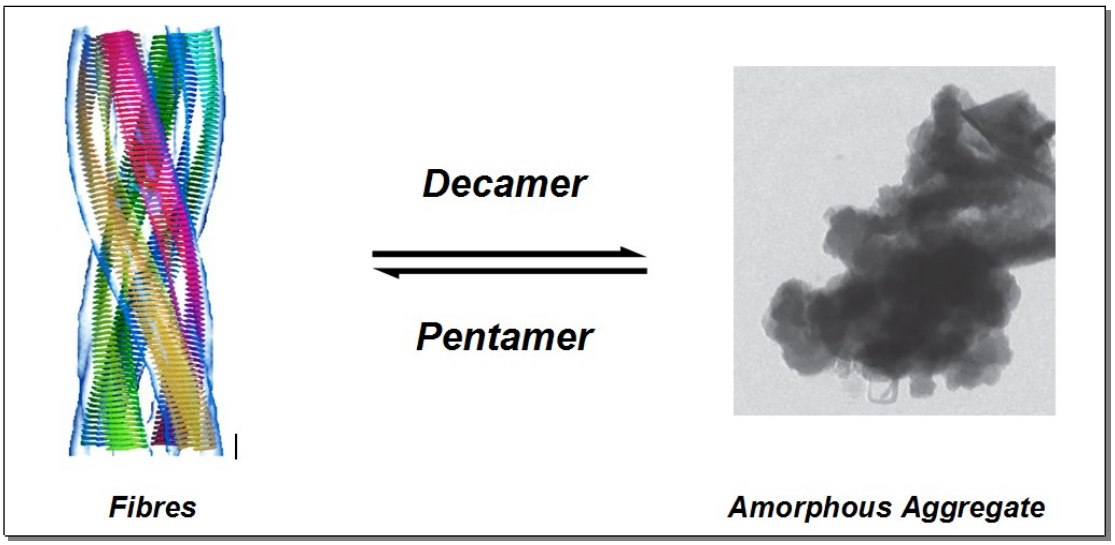


Figure 42: Diagram representing the theory of SAP controlling equilibrium between fibre and amorphous aggregate. The decamers are responsible for shifting equilibrium towards amorphous aggregate and the pentamers are responsible for shifting equilibrium towards cross β-sheets. It may be possible that amyloidosis is just the result of improper regulation of this reversible reaction.

The control THS fibres experienced a similar result to the TE fibres. Those grown in the absence of SAP had a direct correlation between protein precipitation and ThT signal with approximately 73% of the insulin was made into cross β-sheets, 22% was amorphous aggregate and the remaining 5% continued as native insulin or soluble intermediates. Again, as before, those fibres grown in the presence of SAP decamers did not follow this pattern. They had a lower ThT signal yet contained almost equal amounts of precipitate as their SAP free counterparts. Only 37% of the insulin became cross β-sheet, 48% amorphous aggregate with the remaining 15% continuing as either native insulin or soluble intermediates. These results are almost identical to the TE fibres indicating that it is the calcium and not the high salt concentration which is the main effecter of fibre formation in the TC assays.

Fibre Disassembly

Upon the addition of 0.2mg/ml SAP decamers to pre-formed insulin fibres, ThT fluorescence shows that fibre amounts decrease to approximately 50% of their original mass over a period of 3-4 days. Following this initial decline, fibril concentrations will rise again although original fibre amounts are not recovered. It has already been established in this chapter that SAP decamers appear to behave in a fashion similar to chaperonins. They prevent elongation of the fibres by binding to the soluble intermediates and possibly aiding in their refolding into the native state.

Continuing with this concept, one could argue that the reason fibre amounts cease decreasing after the initial 3-4 days is due to saturation of all the SAP molecules. Had the interactions between decamer and intermediate been irreversible then one would expect fibre amounts to remain at equilibrium near this lower fibre mass. The subsequent *increase* in fibre amounts suggests, instead, that the interactions are indeed reversible and that when released back into the solution the protein is able to reassemble itself onto the fibre structure. Whether the protein has gone through the process of re-folding then unfolding again, however, cannot be established as of yet.

The addition of further 0.2mg aliquots of SAP at 5 day intervals again sees a decrease in fluorescence after each addition but as previously experienced, fibre amounts eventually increase even when SAP is in excess. This would suggest that refolding of the insulin into its native state is not occurring. The reason for this is likely to be the lack of ATP and co-chaperones in these simplistic *in vitro* conditions. SAP is unable to undergo the appropriate conformational changes necessary to assist in intermediate folding. Alternatively, it is possible that refolding *is* occurring but the pre-existing fibres are catalysing new fibre growth.

Conclusions

Neutral insulin fibres are successfully made *in vitro* when dialysed stock is subject to continuous inversions on a revolving Stuart Rotator SB3 at room temperature for 10 days. This is confirmed through ThT fluorescence, Congo red birefringence, TEM imagery and CD analysis.

The presence of SAP in these assays, at least in the case of insulin, can have one of two effects depending on its conformation. When in its decameric state, SAP has the ability to reduce total fibre mass. Alternatively, when in its pentameric state SAP has the ability to increase total fibre mass. This would imply that SAP (at least in the case of insulin) is capable of acting as both a molecular chaperone and a fibrillogenic enhancer; it has a dual role. This behaviour of SAP is consistent and is directly proportional to its concentration.

When submit to structural techniques such as TEM and CD, it becomes apparent that the cross β-sheet fibre morphology can vary depending on the conditions in which it was developed. The most stable form is the typical rope-like structure approximately 7-10nm in width and of indefinite length produced in TE buffer conditions. The less stable ‘sickly’ and ‘slate like’ morphologies are experienced in conditions of high salt and calcium respectively. It is not clear as to how salt and calcium disrupt fibre formation but one can assume that they become integrated into the fibre structure and thus disrupt optimum binding of the insulin molecules. Upon inclusion of SAP to these conditions, the more stable rope-like morphology is resumed. This emphasises SAPs involvement in fibre stability.

In addition to these results, another relevant fact was discovered. Cross β-sheet fibres are not the sole product of *in vitro* fibrillogenesis. Amorphous aggregate is also produced. The amount of amorphous aggregate produced again varies between samples of different conditions however a pattern can be determined. When SAP is not involved, there is a direct correlation between protein concentration and cross β-sheet fibre production.

Approximately 80% of the protein is converted to fibres and 16% into amorphous aggregate (although his figure is considerably lower in the TC fibres - 21% ad 69% respectively due to the interference of calcium). Upon the inclusion of SAP decamers there is an increase in amorphous aggregate production and a decrease in cross β-sheet fibre production. This suggests that either the SAP decamers promote amorphous aggregate formation or they promote refolding of unfolded aggregates. The latter is unsuccessful in simplistic *in vitro* conditions, presumably, because of the lack of ATP and co-chaperones. Upon the inclusion of SAP pentamers the reverse is seen. There is a decrease in amorphous aggregate production and an increase in cross β-sheet fibre production supporting the theory that SAP pentamers are promoting the formation of amyloid fibres.

Perhaps SAP is the regulator of a reversible pathway between cross β-sheet fibres and amorphous aggregate; it maintains equilibrium between the two. The decamers are responsible for shifting equilibrium towards amorphous aggregate and the pentamers are responsible for shifting equilibrium towards cross β-sheets. The formation of amyloid diseases may simply be the result of improper regulation of this reversible reaction.

It is plausible that the cross β-sheet formation is a temporary storage solution created in the situation where an excess of unfolded intermediates are present in the serum. By encouraging fibre formation instead of amorphous aggregate formation or soluble oligomerisation, SAP may be preventing damage to the host. One must remember that in small quantities, cross β-sheet folding is not harmful. Only when present in large quantities do cross β-sheet fibres become problematic as they prevent organ function due to stiffening of the tissue. In comparison, soluble oligomers have been known to be toxic *e.g.* Aβ oligomers in Alzheimer's disease

In fact, it is even reasonable to suggest that the role of SAP is to maintain equilibrium between cross β-sheet fibres and the *native* protein. The body has no use for amorphous aggregate so would not wish to produce it deliberately. One could argue against this by highlighting the fact that vast amounts of

amorphous aggregates were produced in these *in vitro* experiments, not native protein. Although this is true, amorphous aggregate was regularly produced in these *in vitro* experiments, this was likely to be the product of the absence of ATP and co-chaperones. Had these been present, the native fold may indeed have been reconstituted upon interaction with SAP decamers.

Chapter 3

Alternate Fibre Analysis

Transthyretin (TTR) Fibre Analysis

Chapter Aims

Following the success of the experiments involving insulin fibres, it was decided to carry out similar assays on fibres made form other proteins. This was to ensure that the dual behaviour of SAP was in fact universal to all proteins and not just unique to insulin. By carrying out the above on a minimum of two more proteins it may be possible for a general 'rule' on SAP behaviour to be established.

Again, it was vital that all fibre assays be performed at neutral pH rather than the acidic conditions normally implemented for fibre formation due to the inclusion of SAP. This was not only to prevent alterations in SAP structure and function, but to make sure that any results obtained could be directly compared those obtained in chapter 2.

In order to monitor SAPs behaviour with respect to another protein the following steps were proposed:

- Produce alternate protein fibres *in vitro* at neutral pH using the rotator method established in chapter 2
- Expose one set of fibres to SAP pentamers during growth.
- Expose another set of fibres to SAP decamers during growth.
- Leave one final set to grow without SAP interference to act as a control.
- Use dyes, electron microscopy and other analytical techniques in order to observe the effect of SAP on fibril formation and morphology.

Transthyretin Structure and Function

The second protein chosen for SAP analysis was transthyretin (TTR). TTR - formerly known as prealbumin - is another protein known to form fibres *in vitro* but, as with insulin, only under acidic conditions.

TTR is responsible for binding and transporting thyroxine and retinol (vitamin A) around the body and has gained its name in doing so - **trans**ports **thy**roxine and **retin**ol. In fact, TTR functions in conjunction with two other thyroid hormone-binding proteins in the serum: thyroxine-binding globulin and albumin. Of these, thyroxine-binding globulin is present at the lowest concentration but has the highest affinity. Albumin is the reverse and is present at the highest concentration of the three but has the lowest affinity.

Unlike albumin, which has a half-life of approximately 20 days, the turnover rate for TTR is rapid with a half-life of only 2–3 days. It is produced in the liver, secreted into the serum and finally catabolized in the kidneys. It has been noted that small quantities of TTR are also synthesised in the choroid plexus and retinal pigment epithelium and is secreted into the cerebrospinal fluid and eye respectively (Murrell et al. 1992).

In the serum TTR exists predominantly as a homotetramer with each of its four identical monomers being 127 residues in length. Each monomer is structured into eight beta strands (A-H) and folded into a β-sandwich. When scrutinised further, one can see that the sandwich consists of two, four stranded beta sheets (CBEF and DAGH) and is held together by loops and a nine residue α-helix connecting strands E and F **(Figure 43)**.

In order for the homotetramer to form, initially two monomers must interact with one another. They bind via hydrophobic interactions which occur between the two F and H strands of neighboring monomers. The final tetramer structure is 55kDa in size, is very stable, and is produced when two of these dimers interact face to face (Ciszak, Cody, and Luft 1992).

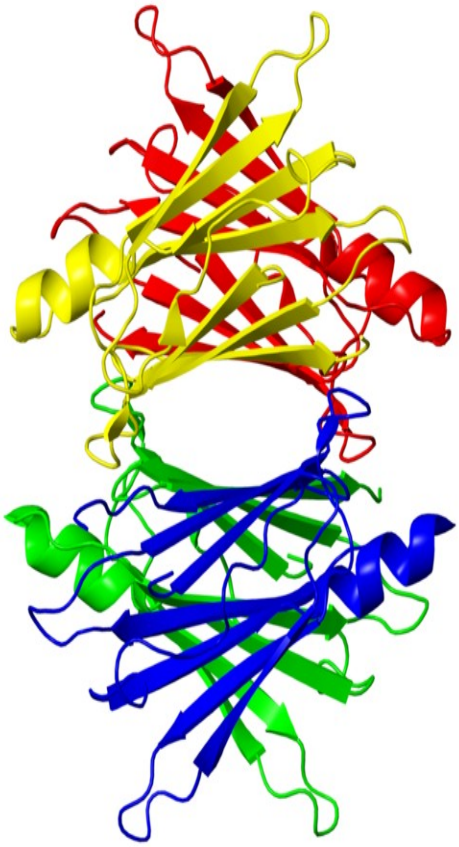


Figure 43: Ribbon diagram of the stable TTR homotetramer. Each monomer is approximately 14kDa in mass and is represented in a different colour.

Transthyretin Amyloidosis

Transthyretin amyloidosis (ATTR) is a term used to describe a group of autosomal, dominant diseases caused by mutated TTR. With the exception of senile systemic amyloidosis (SSA) - which develops from wild type TTR accumulation - all ATTR mutations are inherited, however, they are not found equally worldwide. Instead they are found in clusters within Portugal, Japan and northern Sweden with a gene frequency of 1.5% (Zaros et al. 2008) (M. Olsson et al. 2014). It must also be noted that although present in the body from birth, the symptoms of mutated TTR do not occur until the 3rd, 4th or 5th decade of a patients life. This is because the variant TTR requires several decades to build-up in mass before it becomes problematic (Planté-Bordeneuve and Said 2011) (Pettersson and Konttinen 2010).

As with insulin, TTR fibre formation requires the disassociation of the homotetramer into monomers – this is the rate limiting step. Upon completion of this, the monomers are able to undergo the significant conformational changes required for amyloid fibre formation (Hammarström et al. 2003). Wild type TTR is usually an extremely stable structure, however, when a mutation is present the stability of the TTR homotetramer is compromised leading to the dissociation of the quaternary structure. Sometimes only a single monomer of the quaternary structure carries a mutation where as in other occasions two, three or even all four monomers may carry it.

Mutations in TTR are relatively rare affecting only 1 in 100,000 cases of amyloid, however, TTR mutations must not be treated lightly as they induce a wide set of symptoms and are fatal (Planté-Bordeneuve and Said 2011) (Pettersson and Konttinen 2010).

ATTR is a form of systemic amyloidosis. This means that symptoms are not limited to a single organ but instead affect various organs simultaneously. These include the heart, kidneys, gastro intestinal tract, peripheral autonomic nerves, connective tissue of the carpal ligament and the eyes. The reason

that such a vast array of conditions are experienced with TTR amyloidosis is due to the fact that ATTR is caused by point mutagenesis which is not limited to a single specific point. Alterations can occur at various points along its sequence.

The TTR gene is located on chromosome 18q12.1. More than 120 mutations have been found on this gene, nearly all of which are amyloidogenic. This includes single mutations, compound heterozygotes and deletions. These all lead to a collection of different symptoms which are uniquely categorised into different categories of ATTR. The first and most common of these is when valine at residue 30 is replaced by a methionine forming something more commonly referred to as familial amyloid polyneuropathy (FAP) (Coelho 1996). Alternatively, familial amyloid cardiomyopathy (FAC) is the name chosen when valine 122 is replaced by isoleucine (Jacobson et al. 1997) and other examples include Familial euthyroid hyperthyroxinemia and Senile systemic amyloidosis (SSA).

The most common mutation, Val30Met *a.k.a.* FAP, is responsible for 85% of ATTR cases. It is known to affect the peripheral autonomic nerves, connective tissue of the carpal ligament, heart, kidneys, gastrointestinal tract and eyes. (Tantau, Laszlo, and Laszlo 2015). In FAP, initially it is only the lower limbs that are affected but over time this spreads to the thighs, upper limbs, forearms and fingers as the anterior trunk is involved. Symptoms include discomfort, numbness and spontaneous pain; similar to carpal tunnel syndrome (Ando et al. 2013). A large number of patients with TTR-FAP are often mis-diagnosed with carpal tunnel syndrome and only when a lack of improvement is noticed following carpel tunnel syndrome surgery does correct diagnosis occur (Planté-Bordeneuve and Said 2011).

ATTR is notorious as it frequently affects the heart and is often fatal as a result of this. To date, a total of 14 TTR mutations have been associated specifically with cardiac-predominant familial disease. Nine of these develop an exclusive cardiac phenotype whereas five TTR mutations (Gly30, Arg53, Ser64, His69, Cys114) have also been associated with the nervous system,

manifesting as dementia or cerebral haemorrhage (Tantau, Laszlo, and Laszlo 2015).

Examples of mutations involved in cardiac amyloidosis are Val122Ile, Thr60Ala, Ile68Leu, Leu111Met and Ser77Tyr. The Val122Ile mutation in particular is associated with a late-onset amyloid cardiomyopathy characterized by progressive severe heart failure (Moreira et al. 2010). The consequences of these mutations include increased thickness of both ventricular walls and the interventricular and atrial septums. Deposition of the amyloid on the heart and specifically in the electrical conduction system ultimately results in unpredictable arrhythmia episodes and/or severe conduction disorders. These include high degree atrioventricular blocks with faintness, or even sudden death (Banypersad et al. 2012) (Falk 2005).

Diagnosis and Treatment of ATTR

Diagnosis of both cardiac and neurodegenerative related ATTR is difficult due to the vast array of symptoms. In fact, diagnosis of any form of TTR related amyloidosis is difficult to diagnose due to the extensive collection of diseases related to this particular protein. Be that as it may, a few methods are being used in clinics today.

The first non-invasive test of choice for cardiac amyloidosis diagnosis is echocardiography, or more simply put, a sonogram of the heart. The increased ventricular wall thickness and diastolic dysfunction is one of the earliest echocardiographic found abnormalities. Another test frequently used is cardiac MRI. This provides important information by illustrating the restrictive cardiomyopathy and wall hyperechogenicity (O'Hara and Falk 2003). Finally, Congo red staining can be used alongside iodine^{123} labelled SAP and TTR gene sequencing as a way to establish TTR amyloidosis.

Following ATTR diagnosis the life expectancy for patients is approximately 10 years following symptom onset. To help prolong this, liver transplantation has become the standard treatment form especially in the case of patients with FAP. Transplantation replaces the variant TTR gene with the wild type gene in the liver and causes variant TTR serum concentration to decrease reaching almost zero. However, following transplantation the rate of survival at 5 years is only 59% for non Val30Met patients and 82% for Val30Met patients (Nelson et al. 2013). The disease may progress particularly in patients with mutations which are non-Val30Met. This is as a result of the deposition of wild-type TTR fibrils on pre-existing amyloid matrix. Combined heart and liver transplant is therefore proposed on highly selected patients.

Recently a new method has been developed to help control FAP in particular. Tafamidis and diflunisal are two drugs that are undergoing clinical development worldwide. Tafamidis is a novel TTR stabilizer and diflunisal is a nonsteroidal anti-inflammatory drug that can stabilize TTR tetramers. Suhr *et al.* conducted an 18 month trial which evaluated the status of TTR-FAP patients treated with tafmidis. They used the patients BMI as a way to monitor overall health and improvement. Following the 18 months they observed that the BMI improved in the tafamidis group and worsened in the placebo group. The placebo group was subsequently treated with tafmidis and their BMIs also increased (Suhr et al. 2014). The drug generally caused no harm to the patients and little progression was observed in amyloid cardiomyopathy during the course of treatment. However, these drugs do have gastrointestinal side effects and vaginal and urinary tract infections may appear.

TTR Fibre Formation *in Vitro*

Wild type TTR has shown to be extremely resistant to aggregation at physiological pH due to the tetramer being extremely stable in these conditions (Bourgault et al. 2011); the binding of a single ligand alone is enough to completely prevent amyloidosis. Various lab groups have demonstrated this *in vitro*. They have shown that in addition to thyroxine and

retinol, TTR has the ability to bind numerous other small molecules in its thyroxine binding sites. These include many natural products (such as resveratrol), drugs (Tafamidis or Vyndaqel, diflunisal, flufenamic acid), and toxicants (polychlorinated biphenyl - PCB). Upon binding of these molecules, TTR is stabilised and unable to form amyloid fibres (Razavi et al. 2003) (Adamski-Werner et al. 2004) (Purkey et al. 2004) (Baures et al. 1999).

Studies using acidic conditions have shown that a lower pH can be used to facilitate homotetramer dissociation, however there have been cases where pH's as low as 4.6 have still not induced alterations in TTR structure. Significant structural changes are only noted when pH is lowered to 4 or less. Further to this, in 1992 Colon and Kelly found that the optimum range for producing TTR fibres is between pH 3.6 and 4.8. Above this range the stability of the homotetramer is too strong and below this the rate of fibril formation decreases markedly.

During their investigations Colon and Kelly also discovered some information regarding the structure of their TTR fibres. Using Far and Near-UV CD methods they were able to demonstrate that under acidic conditions TTR has the ability to retain the majority of its tertiary form, even when part of the cross-β sheet. In doing so they provided further evidence that only partial denaturation is required in order for fibres to form rather than complete unfolding. Some years later the cause of the unfolding was presented by another group. In 2008 Palaninathan *et al.* discovered that under acidic conditions the EF helix (residues 75-90), the adjacent EF loop and part of the beta-strands become unstructured allowing TTR fibres to form.

Unfortunately, in order for any TTR results to be comparable to those produced with insulin in chapter 2 it was essential that fibres be created at neutral pH, not acidic pH, preferably using the rotator method. Due to the high stability of wild type TTR at physiological pH (and even at acidic pH!) it was deemed unlikely that fibres would form using this method and any attempt to do so would be futile. However, as a way to combat this, it was decided to use less stable, mutant variants.

L55P and S52P are two different TTR variants known to dissociate into monomers under physiological conditions at 37°C as shown by Lashuel *et al*. in 1999. In the L55P variants a proline residue has replaced the wild type leucine at position 55 whereas in the S52P proline residue has replaced the wild type serine at position 52. These mutations cause destabilisation of the D-strand on the outer edge of the tetramer and result in the formation of a disordered loop. The disordering of this loop facilitates tetramer dissociation and thus provides monomers capable of unfolding into amyloid prone intermediates.

Methods

TTR Preparation

Two different TTR mutants were used during the course of experimentation: L55P and S52P. Each mutant was expressed and purified by a predecessor of the Crystallography Lab at the Royal Free Hospital. Stock solutions were stored in 25mM TRIS, 300mM NaCl buffer solution at a temperature of 4°C.

TTR Fibre Formation

L55P and S52P TTR stocks were diluted to 1ml samples of 1mg/ml. This was achieved using TE buffer in a 1.5ml eppendorf. The samples were continuously inverted – or "flipped" - at a speed of 20rpm on a revolving Stuart Rotator SB3 for 10 days at room temperature as per the protocol for making insulin fibres described in chapter 2.

Sample	TTR (5 mg/ml)	TE Buffer	dH_2O	Total TTR Concentration
1a	20 µl	800 µl	180 µl	0.1 mg/ml

To monitor the effect of protein concentration of fibre formation, TTR stocks were diluted to 1ml samples of various concentrations ranging from 0.1mg/ml - 1mg/ml. This was achieved using TE buffer in a 1.5ml eppendorf. The samples were continuously inverted – or "flipped" - at a speed of 20rpm on a revolving Stuart Rotator SB3 for 30 days at room temperature as per the protocol for making insulin fibres described in chapter 2. This protocol was repeated for both the L55P and S52P variants.

Set up:

Sample	TTR (5 mg/ml)	TE Buffer	dH_2O	Total TTR Concentration
1b	20 µl	800 µl	180 µl	0.1 mg/ml
1c	40 µl	800 µl	160 µl	0.2 mg/ml
1d	100 µl	800 µl	100 µl	0.5 mg/ml
1e	200 µl	800 µl	-	1mg/ml

Fibre Formation with SAP decamers

To observe the effect of SAP decamers on TTR fibre growth, the "TTR Fibre Formation" method was repeated with the addition of 0.2mg/ml SAP. This provided a stoichiometric ratio of 1:10 of SAP decamers to TTR respectively. Each sample was continuously inverted at a speed of 20rpm on a revolving Stuart Rotator SB3 for 30 days at room temperature. This protocol was repeated for both the L55P and S52P variants.

Sample	TTR (5 mg/ml)	TE Buffer	SAP (6.66mg/ml)	dH_2O	Total TTR Concentration
2a	20 µl	770 µl	-	210 µl	0.1 mg/ml
2b	20 µl	770 µl	30 µl	180 µl	0.1 mg/ml
2c	40 µl	770 µl	-	190 µl	0.2 mg/ml
2d	40 µl	770 µl	30 µl	160 µl	0.2 mg/ml
2e	100 µl	770 µl	-	130 µl	0.5 mg/ml
2f	100 µl	770 µl	30 µl	100 µl	0.5 mg/ml
2g	200 µl	770 µl	-	30 µl	1mg/ml
2h	200 µl	770 µl	30 µl	-	1mg/ml

Fibre Formation with SAP Pentamers

To perform this particular assay the SAP was diluted to 0.2mg/ml using TC buffer. This was left to incubate for half an hour to ensure saturation of SAP calcium pockets. Following this, various concentrations of TTR were added and the samples were continuously inverted at a speed of 20rpm on a

revolving Stuart Rotator SB3 for 30 days at room temperature. This protocol was repeated for both the L55P and S52P variants.

Sample	TTR (5 mg/ml)	TC Buffer	SAP (6.66mg/ml)	dH_2O	Total TTR Concentration
3a	20 µl	770 µl	-	210 µl	0.1 mg/ml
3b	20 µl	770 µl	30 µl	180 µl	0.1 mg/ml
3c	40 µl	770 µl	-	190 µl	0.2 mg/ml
3d	40 µl	770 µl	30 µl	160 µl	0.2 mg/ml
3e	100 µl	770 µl	-	130 µl	0.5 mg/ml
3f	100 µl	770 µl	30 µl	100 µl	0.5 mg/ml
3g	200 µl	770 µl	-	30 µl	1mg/ml
3h	200 µl	770 µl	30 µl	-	1mg/ml

Control

To see the effect the salt was having on TTR fibre formation the rotator assays were repeated in a High Salt buffer ("THS Buffer" - 20mM TRIS, 600mM NaCl, 0.01% Na Azide, pH8). A range of TTR concentrations spanning 0.1-1mg/ml of were continuously inverted at a speed of 20rpm on a revolving Stuart Rotator SB3 for 30 days at room temperature in the presence and absence of 0.2mg SAP. This protocol was repeated for both the L55P and S52P variants.

Sample	TTR (5 mg/ml)	THS Buffer	SAP (6.66mg/ml)	dH_2O	Total TTR Concentration
4a	20 µl	770 µl	-	210 µl	0.1 mg/ml
4b	20 µl	770 µl	30 µl	180 µl	0.1 mg/ml
4c	40 µl	770 µl	-	190 µl	0.2 mg/ml
4d	40 µl	770 µl	30 µl	160 µl	0.2 mg/ml
4e	100 µl	770 µl	-	130 µl	0.5 mg/ml
4f	100 µl	770 µl	30 µl	100 µl	0.5 mg/ml
4g	200 µl	770 µl	-	30 µl	1mg/ml
4h	200 µl	770 µl	30 µl	-	1mg/ml

TTR Fibre Analysis using ThT

Please refer to the 'Fibre Analysis using ThT' section in chapter 2.

Fibre Detection using Congo red

Please refer to the ‘Fibre Detection using Congo red’ section in chapter 2.

Fibre Analysis using Circular Dichroism

Please refer to the ‘Fibre Analysis using Circular Dichroism’ section in chapter 2.

Fibre Formation in Alternate Conditions

Fibre Formation via Seeding

Each variant of TTR was “seeded” in an attempt to catalyse fibre growth. This was attempted two ways:

The first way involved repetition of the “TTR Fibre Formation” method with an additional 30 µl (0.03mg) of either pre-aggregated TTR or insulin added to the new assay. This mixture was subsequently rotated for 10days at 20rpm on a revolving Stuart Rotator SB3.

The second method involved using sodium acetate buffer at pH5 and 6 (50mM sodium acetate, 100mM potassium chloride) as taken form the Takahiro Tajiri (2002) protocol. Here, 0.2mg/ml off TTR was incubated in this buffer for 10 days at 37°. This, again, was attempted with and without 30 µl (0.03mg) TTR aggregate and Insulin fibre “seeds”

Fibre Formation – Tajiri Protocol (Takahiro Tajiri 2002)

Each variant of TTR was diluted to the following range of concentrations: 0.1mg/ml, 0.2mg/ml, 0.5mg/ml and 1mg/ml. This was achieved using sodium acetate buffer (50mM sodium acetate, 100mM potassium chloride, pH5) and incubated at 37°C for 10 days. This was later repeated at pH 6 and 7 with and without the addition of 0.2mg/ml SAP

Sample	TTR (5 mg/ ml)	Takahiro Buffer	SAP (6.66mg/ ml)	dH_2O	Total TTR Concentration
5a	20 µl	770 µl	-	210 µl	0.1mg/ml
5b	20 µl	770 µl	30 µl	180 µl	0.1 mg/ml
5c	40 µl	770 µl	-	190 µl	0.2 mg/ml
5d	40 µl	770 µl	30 µl	160 µl	0.2 mg/ml
5e	100 µl	770 µl	-	130 µl	0.5 mg/ml
5f	100 µl	770 µl	30 µl	100 µl	0.5 mg/ml
5g	200 µl	770 µl	-	30 µl	1 mg/ml
5h	200 µl	770 µl	30 µl	-	1 mg/ml

Fibre Formation – Kugimiya Protocol (Kugimiya et al. 2011)

A total of 1.75mg/ml of each TTR strain was incubated at 37°C in sodium acetate buffer for 5 days in the dark (20 mM sodium acetate and 100 mM NaCl at pH 3.0). SAP was not added to this assay due to the highly acidic conditions which would likely alter its structure and function.

Sample	TTR (17.5 mg/ml)	PBS Buffer	Total TTR Concentration
6a	100 µl	1400 µl	1.2mg/ml

Results and Discussion

TTR Fibrillogenesis (Rotator method)

L55P and S52P TTR stocks were diluted to concentrations of 0.2mg/ml and made up to 1ml using TE buffer (pH 8). The samples were continuously inverted – or "flipped" - at a speed of 20rpm on a revolving Stuart Rotator SB3 for 10 days at room temperature and the results analysed via ThT fluorescence.

Within the first 3 or 4 days of inversion, as per the insulin assays, white precipitates developed in both the L55P and S52P TTR variant samples. This precipitation increased in density over the course of 2-3 days and then appeared to remain constant for the rest of the protocol. The total amount of precipitant produced was considerably lower to what had previously been seen. Only approximately $^1/_6$ of precipitate was produced compared to the insulin assays. On the tenth day the TTR samples underwent ThT analysis in order to determine whether the precipitate was cross β-sheet in nature or amorphous aggregate. Unfortunately, in the case of TTR, neither the L55P nor the S52P variant showed enhanced fluorescence between wavelengths 420nm and 600nm on a spectrophotometer. This suggests that the rotator method had only produced amorphous aggregate and not cross β-sheet fibres. This was verified using Congo red analysis. Here the samples showed no apple green birefringence under polarised light.

This was an unexpected result. The rotator method had proven to be overwhelmingly successful in the case of the protein insulin. It may be true that wild type TTR is extremely resistant to aggregation at physiological pH due to tetramer stability, however, both mutants (L55P and S52P) are known for their high tendency to dissociate into monomers *in vitro* as discovered by Lashuel *et al.*

Further investigation of the literature suggested various explanations as to why cross β-sheets had not been produced. Lashuel explains that high TTR concentrations can restrict fibre formation. Lashuel also indicates that aged L55P TTR, in particular, has a tendency to form soluble aggregates which can interfere with experiments. Upon discovery of these facts, all TTR stocks were treated as suggested in this paper. They were subject to centrifugation at 50,000 rpm followed by filtration through a 0.2μm membrane so to remove the vast majority of soluble aggregates. Following this treatment the rotator protocol was repeated with a few alterations in an attempt to encourage fibre growth. These alterations included sample concentration, rotator speeds, temperature, protocol duration, buffer conditions and SAP incorporation.

Samples of each variant were diluted to varying protein concentrations ranging from 0.1-1mg/mg. This was done in all three buffer types previously used for insulin fibre formation - TE buffer, TC buffer and THS buffer – in the presence and absence of SAP. These were all submitted to the rotator method using various speeds of 2-30 rpm, (the minimum and maximum speed respectively) for a longer duration of one month. Replicas of these assays were also subject to incubation at 37°C with/without agitation. Sadly, none of these methods produced cross β-sheet fibres. All samples developed white precipitates but all were amorphous in nature according to ThT and Congo red analysis.

Structural analysis using the CD technique supported these findings **(Figure 38)**. The spectrum shows a curve of TTR aggregate compared to its insulin fibre counterparts. This particular TTR sample demonstrated in the spectrum was produced from the inversion of 0.2mg/ml L55P TTR in TE buffer for one month - The other samples from both variants produced almost identical curves but were excluded from the graph for display purposes. The curve produced was typical of neither α-helices nor β-strands. This was true of all samples produced from both TTR variants. They each produced a very broad curve just above the base line with no x axis interception point. This indicates that the aggregates formed using the TTR variants contained no secondary structure and was indeed amorphous in nature.

What this suggests is that the conditions used on the TTR variants were sufficient to allow denaturation of the TTR molecules but not optimal to allow cross β-sheet formation. Perhaps the TTR molecules were not in their monomeric form when aggregation occurred thus would not assemble into cross β-sheets. Alternatively it is possible that cross β-sheet fibres *were* being made in conjunction with amorphous aggregate but just at very low concentrations. The insulin assays from chapter 2 clearly demonstrate the ability of proteins to make both cross β-sheet fibres and amorphous aggregate simultaneously. They have also shown that the ratio of fibres and aggregate varies depending on the conditions in which they were made. It is entirely possible that TTR, being a different protein to insulin, would behave differently in these assay conditions and thus produce a different ratio of cross β-sheet fibres to amorphous aggregate. It is possible that cross β-sheet fibres *are* present in the assays but just in very small quantities.

Upon further scrutiny of the TTR results using the ThT method, it was noticed that a negligible 'blip' of ThT fluorescence can be seen. The spectrophotometer shows enhanced fluorescence between wavelengths 420nm and 600nm with a peak of approximately 2 at 476.5nm **(Figure 44)**. At first this reading was considered anomalous due to its small size, however, every sample produced this 'blip' where as native TTR did not. Although this was a very small, but promising, development, the presence of cross β-sheet fibres still had to be confirmed using an alternative method so to confirm this was not a false positive result.

When magnified to x1600 on a Zenith Microlab ML-1000M Monocular Laboratory Microscope 60040, the Congo red slides of each TTR assay confirmed very small amounts of fibres. All samples showed specks of the apple green birefringence typical of cross-β sheets under polarised light. This would suggest that, in very small amounts, TTR cross β-sheet fibres are indeed being formed using the rotator method at neutral pH in all three buffer types. This is true of both mutants and can be replicated **(Figure 45)**.

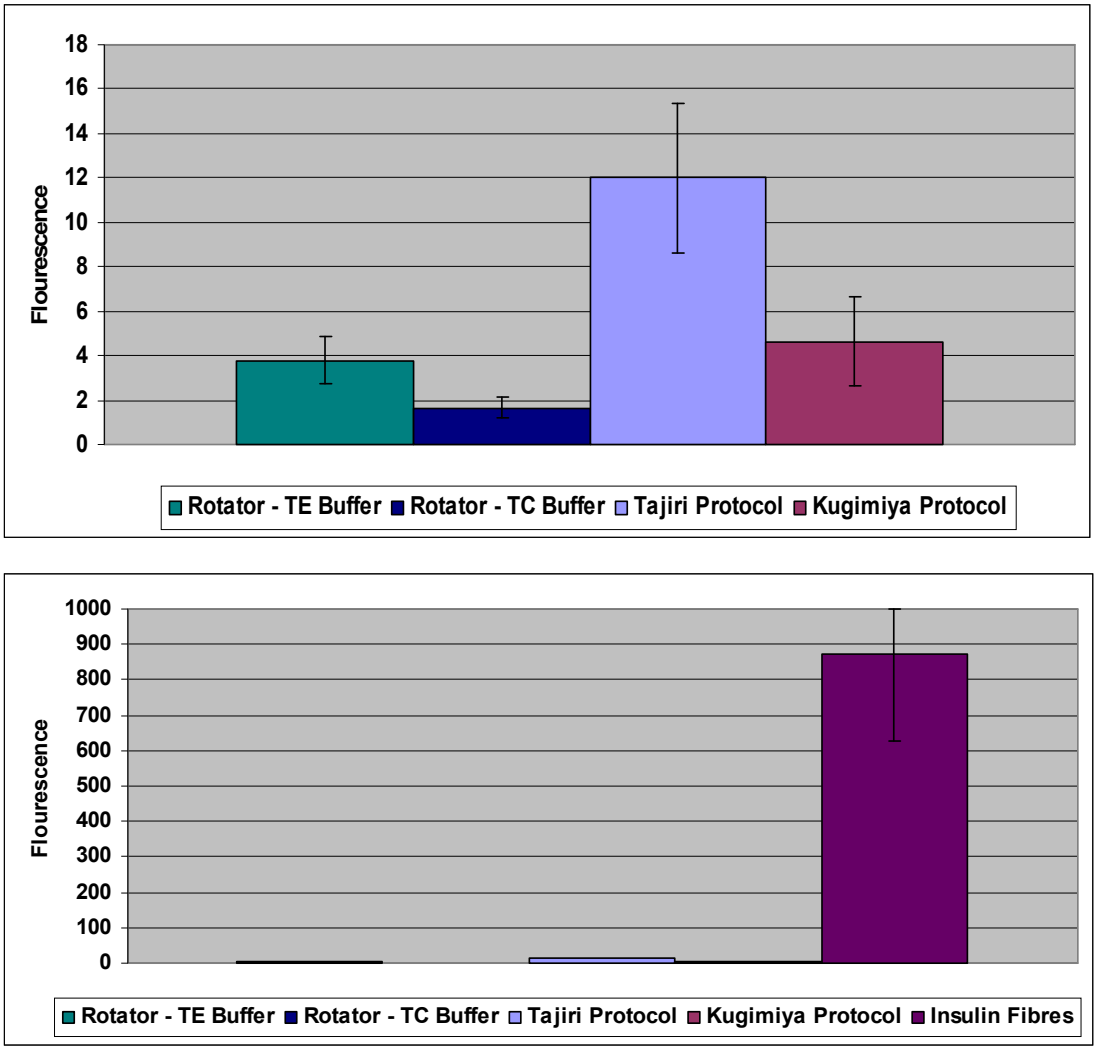


Figure 44: (Above) A bar graph comparing average ThT fluorescence of neutral TTR fibres created in TE buffer and TC buffer via the rotator method alongside acidic TTR fibres created via the Tajiri and Kugimiya protocol. Each bar is an average of three samples and experimental error bars are displayed. The fibres grown in TE buffer reached an average peak in fluorescence of 3.8 where as those grown in TC buffer produce an average of 1.6. Those grown in via the Tajiri and the Kugimiya protocols produced fluorescence intensities of 12.0 and 4.6 respectively. (Below) A graph comparing the aforementioned TTR fibres with TE insulin fibres grown using the rotator method at neutral pH. Insulin fibres reach an average fluorescence intensity of 871.2.

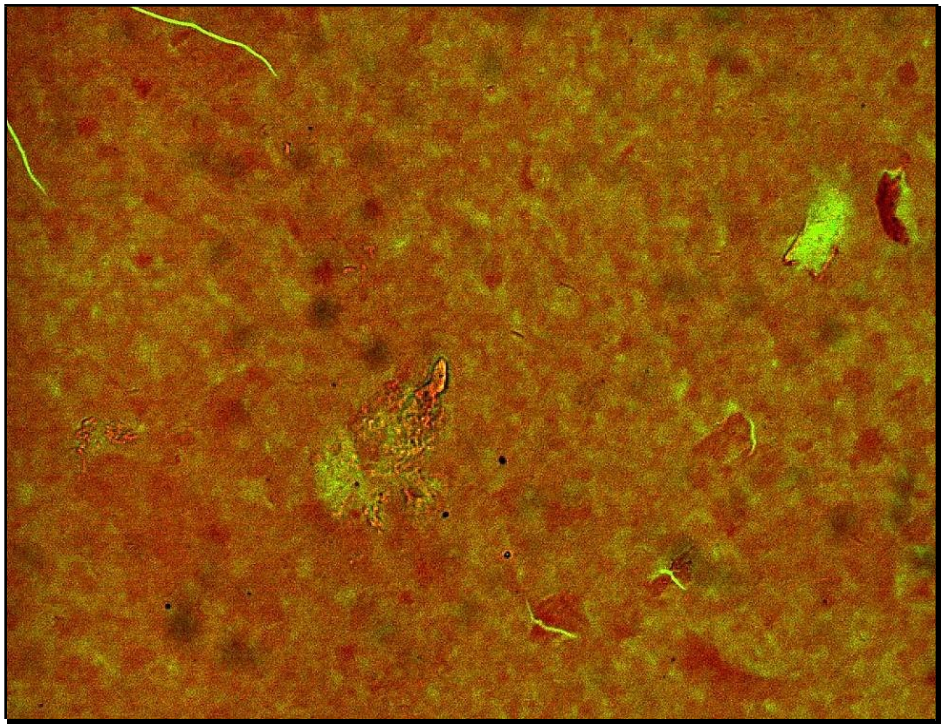


Figure 45: A photograph of TTR fibres stained with Congo red grown in TE buffer via the rotator method. The image was viewed using a Zenith Microlab ML-1000M Monocular Laboratory Microscope 60040 at a magnification of x1000. The sample shows traces of the apple green birefringence typical of cross-β sheets under polarised light.

Unfortunately, due to the miniscule amounts of fibre it is difficult to compare TTR results to those previously obtained by insulin. Also, the different effect (if any) of SAP pentamers and decamers on each mutant is not distinguishable. Due to this, alternate methods of forming TTR fibres had to be carried out.

Fibre Formation via Seeding

'Seeding' was another method used to attempt TTR fibre formation. "Seeding" is a method that is able to catalyse fibre growth. It involves the addition of pre-formed fibre nuclei to a solution of soluble protein. In doing so, the time consuming lag phase often seen in fibre growth is eliminated and rapid fibre formation is achieved.

Seeding is normally performed between variants of the same protein therefore it was attempted to 'seed' fresh TTR assays with 30 µl (0.03mg) of pre-aggregated TTR. The specks of apple green birefringence and increased ThT fluorescence demonstrated that cross β-sheet nuclei had indeed formed. These nuclei have the ability to catalyse cross β-sheet growth *in vitro*. After rotation for 10 days with the 'seed' however there was no improvement on fibre amounts. Again, a white precipitate was formed but it was mostly amorphous aggregate in nature. This result is most likely due to the fact that in the pre-formed TTR fibres, the cross β-sheets were in such small quantities that there was no guarantee that they would be present in the 30 µl sample added to the fresh assay. It would be more beneficial to separate the small amount of TTR fibres from the amorphous aggregate and *then* seed, however, it is impossible to separate the two substances. Both amorphous aggregate and cross β-sheet fibres are white solids. They each have very similar, if not identical, densities and are made of the same protein.

Attempts were made to seed the TTR variants using insulin fibres. It has been shown in literature that cross-species seeding is possible, however, this is usually only successful when both proteins have significant tertiary structural similarity as explained by Ma and Nussinov in 2012. The structure of insulin

has already been described in chapter 2 as being a predominantly helical structure where as TTR exists mostly as β-sheets. This indicates that the two are not compatible with regards to seeding as was demonstrated. This was verified when TTR failed to produce cross β-sheet fibres upon seeding with insulin fibres.

Fibre Formation in Alternate Conditions

Another method implemented in an attempt to increase TTR fibre amounts was with the use of acidic buffers. This was not an ideal experiment with regards to this investigation due to two reasons: The first being that acidic conditions are not necessarily representative of the conditions experienced within the body. The second is that the acidic conditions could affect the interaction of SAP with fibres due to the occurrence of SAP denaturation. It was chosen, however, to attempt acidic fibre formation regardless so to ensure that TTR fibre formation was possible to the same extent as with insulin.

Tajiri *et al.* had found success in forming TTR fibres using sodium acetate buffer in 2002. An attempt was made to replicate this. Each sample went cloudy after three days and was scrutinised using ThT upon completion of the protocol. Again, initially it appeared that this method had also been unsuccessful. Under further scrutiny, however, minuscule amounts of fibre were detected using ThT. Approximately 3 times more TTR fibre was experienced with this protocol than with the rotator method at neutral pH. A ThT peak reading of 12 ±28.2% was achieved. The Kujimiya protocol from 2011 yielded similar success. A peak reading of 5 ±28.2% was achieved – similar to the neutral rotator method in TE buffer. From figure you can see that of all these methods the TC rotator fibres produced the smallest amount of fibres at on 1.65. However, these results are so small and the instrumental error margins so large in comparison that this observation is not justified. Also, when compared to the insulin fibre fluorescence in figure all the TTR results become negligible.

Conclusion

Although wild type TTR is notorious for resisting monomerisation even at very low pH, the L55P and S52P variants are not. Therefore, the difficulty experienced in obtaining fibres in this chapter was unexpected.

It is not completely clear as to where the fault lay in these particular experiments. It is true that the variant TTR samples used during experimentation were a year old, however proper measures were taken to ensure removal of obstructive, solubilised components. Perhaps it was the storage of the TTR variants which led to the lack of results. It may have proven to be more successful to freeze the stocks at lower concentrations in order to prevent oligomerisation form occurring.

When looking at the pre-existing literature one will notice that the many of the papers illustrating TTR fibres made *in vitro* have done so using ThT analytical methods alone. The unreliability of this method has already been stated various times in chapters 1 and 2. ThT is known to undergo enhanced fluorescence following non–specific binding to bacteria, connective tissue and genetic material such as DNA and RNA. Perhaps a percentage of these 'TTR fibres' found in literature are not cross β-sheet fibres at all. False positives could have occurred and there is not enough evidence to dictate otherwise. Not all the papers have used alternative methods to verify the presence of the cross β-sheets so this is entirely plausible.

Also, when looking at fluorescence intensity it is important to remember that the numbers obtained from each spectrophotometer machine cannot be directly compared to those of another machine. The resulting intensity numbers are determined by the software, the detector systems, the electronics and filtering, etc of the machine in question thus making them unique and of no units. This could explain why when carrying out the acidic protocols on the two TTR variants – the protocols known to produce TTR fibre as confirmed in the literature – only a very small ThT peak was observed. It is possible that TTR cross β-sheet fibre amounts obtained in this chapter were

the same as in the literature but the different spectrophotometers are generating different numbers. Perhaps insulin produces such a vast amount of fibres that it is making the TTR results appear insignificant in comparison.

Another reason for taking extreme care when comparing fibres from different papers is that a number of factors are known to affect fluorescence intensity; factors which scientists are not always considerate of or aware of. Different environmental conditions, the specific protein involved, fibril morphology, ThT concentration, pH, and ionic strength can all have an affect on fluorescence. There is no consistency within the literature regarding ThT concentration or environmental conditions in which the TTR 'fibres' were formed. This means that pre-existing literature regarding TTR fibres cannot be compared to one another let alone to the results obtained here. Only by attempting fibrillogenesis with an alternate protein can matters be clarified.

Amyloid Beta (Aβ) Fibre Analysis

Amyloid Beta (Aβ) Structure and Function

Amyloid Beta (Aβ) was the final protein chosen to undergo fibre analysis with SAP. It was chosen due to the fact that it is a natively unfolded protein. By existing in such a form, any issues regarding monomerisation of stable quaternary structures are eliminated. This in turn should make amyloid prone intermediates easy to obtain.

Aβ derives from the cleavage of amyloid precursor protein (APP), a transmembrane protein expressed at high concentrations in the brain and neuron synapses of the nervous system. In order to produce Aβ from APP, the APP molecule is cleaved in its transmembrane region by secretase enzymes β and ɣ. It should be noted that cleavage via the β- and γ-secretases can be promiscuous and often results in several isoforms of Aβ fragments being produced, some of which are more amyloidogenic than others. The protein chain can range between 30 and 51 residues in length with the most common isoforms being $A\beta_{40}$ and $A\beta_{42}$ (F. Olsson et al. 2014)

At present, the physiological role of Aβ is unknown but various theories have been presented. Tabaton *et al.* believe Aβ to be involved in the activation of kinase enzymes, in particular, the ones responsible for its own production; β and γ secretase enzymes. They noticed that Aβ, when in high concentrations, has the ability to behave as an extracellular signalling peptide. It begins a cascade which regulates both β– and γ–secretase activity, thus regulating both steps of its own cleavage from the Aβ precursor protein. This in turn creates more Aβ and begins a positive loop effect on Aβ production (Tabaton et al. 2010). Alternatively, multiple studies have shown Aβ to protect neurons either by shielding them against oxidative stress, or through the regulation of cholesterol transport (Zou et al. 2002) (Yao and Papadopoulos 2002). All the same, a number of animal studies have demonstrated that physiological

function is not disrupted in the absence of Aβ, implying that its role can be considered negligible (Luo et al. 2003).

Aβ Amyloid Development

Under normal conditions, there is equilibrium between the production and elimination of Aβ that maintains Aβ at constant levels. This is referred to as the steady state (Shankar and Walsh 2009). However, in aging and excitotoxicity, the steady sate is disrupted leading to an accumulation of Aβ and senile plaque formation.

An example of the steady state being disrupted can be explained in conjunction with β-Site APP-cleaving enzyme 1 (BACE1) - the major β-secretase in the brain - and γ-secretase. Upon cleavage of APP by BACE1, a neurotoxic form of Aβ is created in the form of the C99 fragment. This is further cleaved by γ-secretase to produce the final Aβ chain (Cheng et al. 2011). Aβ amyloid often occurs when the regulation of the proteolytic function of γ-secretase is compromised resulting in high concentrations of neurotoxic C99 chains. This often happens when the multi-subunit γ-secretase complex is compromised. γ-secretase is made up of four core proteins: presenilin 1 (PSEN1) and 2 (PSEN2), nicastrin and anterior pharynx-defective 1 (APH-1). Mutations in any of these four result in the increased ratio of Aβ in early-onset forms of Alzheimer's disease (Ridge, Ebbert, and Kauwe 2013).

Another way in which Aβ imbalance can be caused is through improper clearance of the protein from the brain. Clearance can happen in a number of different ways. These include with P-glycoprotein and with low-density lipoprotein receptor-related protein 1 (LRP1). In 2005 Cirroto *et al.* demonstrated that removal of P-glycoprotein enhanced deposition in the brain of an Alzheimer's disease mouse model (Cirrito et al. 2005) and more recently in 2013 Kanekiyo *et al.* showed an increase of amyloid plaque deposition in the cortex in LRP1 knock out mice (Kanekiyo et al. 2013).

A final way in which Aβ clearance can be regulated is through the use of degrading peptidase enzymes, in particular, neprilysin enzyme and insulin degrading enzyme. Studies have shown that neprilysin knockout mice have increased levels of Aβ peptides in the brain. Moreover, the activity of neprilysin is reduced in the cortex and hippocampus of Alzheimer's patients (Caccamo et al. 2005). In the case of insulin degrading enzyme, genetic studies have shown that insulin-degrading enzyme gene variations are associated with the clinical symptoms of Alzheimers disease (W. Q. Qiu and Folstein 2006).

Alzheimer's Disease

Although often associated with Alzheimer's, Aβ is not the sole protein responsible for this particular disease. It is found in conjunction with intracellular neurofibrillary tangles containing tau protein. Intracellular neurofibrillary tau tangles are similar to amyloid plaques in that they are formed from cross β-sheet fibres, however, they remain as such and do not associate with extracellular matrix proteins and proteoglycans. Another difference between the Aβ plaques and the tau tangles is their site of deposition. Aβ plaques are found between brain nerve cells (neurons) whereas tau tangles are found within the neurons themselves (Butterfield and Boyd-Kimball 2004) (Ittner et al. 2010). The accumulation of these two proteins causes damage in the neurons resulting in mental and cognitive deficits such as impaired memory, intellect and personality disorder (Khatami 2012). Collectively, Alzhemer's disease related problems reduce quality of life, cause physical disability and ultimately decrease life expectancy (C. Qiu, Kivipelto, and von Strauss 2009) **(Figure 46)**.

Alzheimer's disease is considered to be the most widespread neurodegenerative disorder in the world. It is responsible for 75% of all dementia cases and affects approximately 35.6 million people. This number is believed to be increasing in correlation to population aging and it is estimated that 106.8 million people will be affected by 2050 (Nordberg 2008).

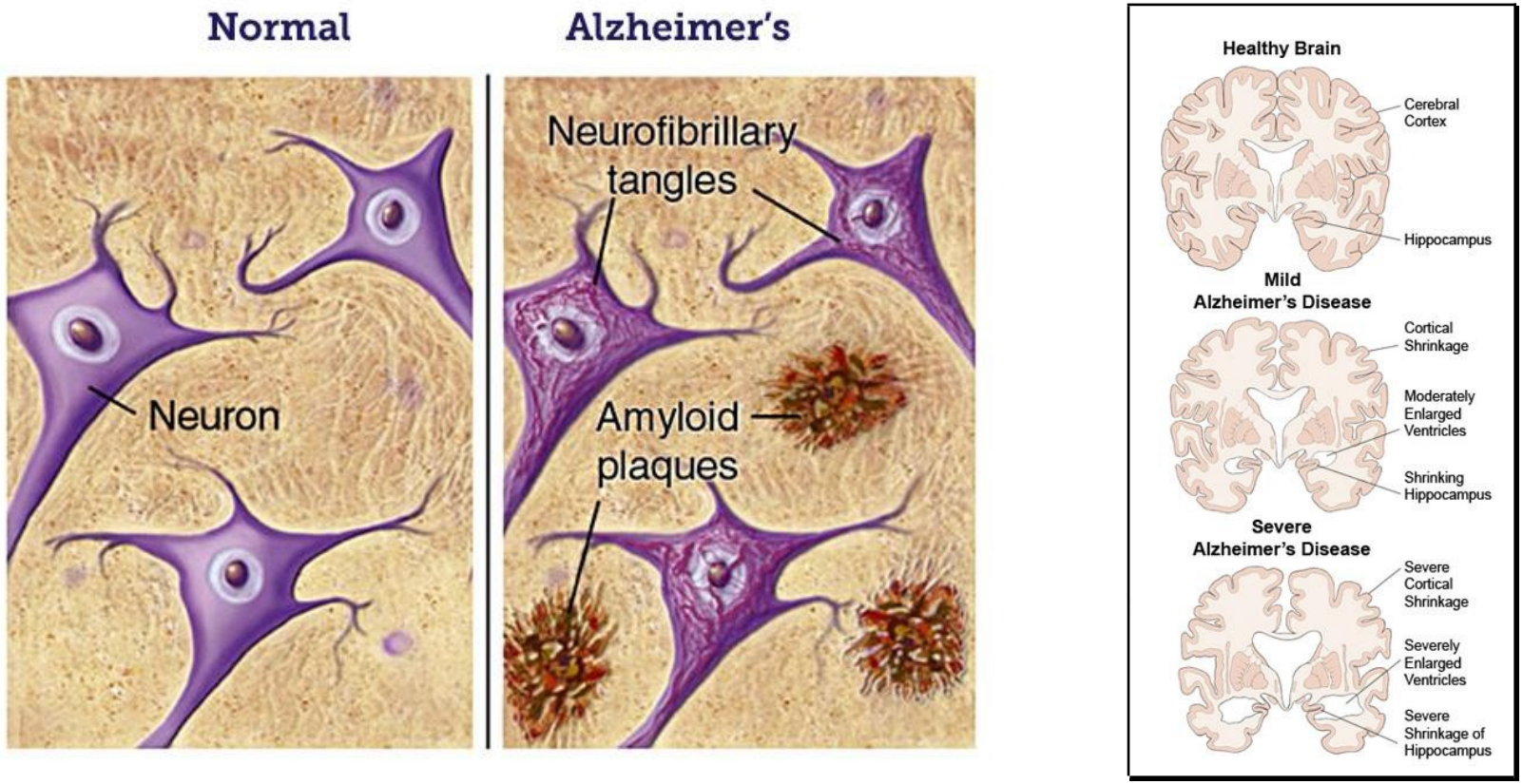


Figure 46: (Left) Comparison of normal neurons with those disrupted with plaques and neurofibrillary tangles. The amyloid plaques are located between the neurons and are composed of Aβ fibrils, extracellular matrix proteins and proteoglycans. The tangles are made from tau fibril alone and are located within the neuron themselves. (Right) Comparison of normal brains with those mildly and severely disrupted from Alzheimer's (Bright Focus, 2015).

Alzheimer's disease can be either inherited (familial) or spontaneous (sporadic). The familial version arises due to mutations within the amyloid precursor protein (APP) gene whereas the sporadic version occurs when an extra copy of the APP gene is produced. Aβ amyloidosis can also be caused by defects in its regulatory enzymes as explained above (Nunomura et al. 2006).

Another point to consider is that in their aggregated form, Aβ peptides are able to induce neurotoxicity (Moreira et al. 2010). Neurotoxicity refers to toxicity in the nervous system. It occurs when exposure to natural or artificial toxic substances (neurotoxins) alters the normal activity of the nervous system in such a way as to cause damage to nervous tissue. Aβ aggregates induce this by provoking nitric oxide formation and an influx of calcium ions. This disrupts synaptic functions and can eventually lead to the formation of peroxynitrite radicals and cell death (Kihara and Shimohama 2004). Unfortunately the precise mechanism of this is still unknown. Interestingly, the accumulation of soluble Aβ oligomers (nonfibrillar), has the same neurotoxic effects (Hardy and Selkoe 2002). This indicates that neurotoxicity is induced through high concentrations of Aβ rather than by the cross β-sheet structure itself.

Treatment

Various drug treatments have been used over the years for Alzheimers disease including cholinesterase inhibitors, memantine and anti-depression/anxiety drugs. These treatments do not slow or stop the progression of the disease, but they may help with the symptoms for a time. The reason for this is that none of these treatments act on the plaques themselves but instead reduce the effect of the chemical changes occurring in the body. For example, the cholinesterase drug acts as an acetylcholinesterase inhibitor thereby inhibiting the acetylcholinesterase enzyme from breaking down acetylcholine, Acetylcholine is the neurotransmitter released by motor neurons of the nervous system in order to activate muscles. By increasing both the level and duration of action of

the neurotransmitter acetylcholine a patient is able to maintain a greater amount of muscular control. Moreover, memantine is a blocker of *N*-methyl-D-aspartate receptors (NMDA receptors). The NMDA receptors are a glutamate receptor and an ion channel protein found in the nerve cells. It is activated when glutamate and glycine bind to it, and when activated it allows positively charged ions to flow through the cell membrane. The NMDA receptor is very important for controlling synaptic plasticity and memory function.

Due to the inability to slow the progression of the disease, remarkable efforts have been made to find diagnostic markers which predict the disease earlier. Neuroimaging methods such as magnetic resonance imaging and positron emission tomography have been developed to enable researchers to diagnose Alzheimer's disease in its early stages (Graham 2011). Also, several biomarkers, which are crucial in detecting pathological features of AD, have been found in cerebrospinal fluid (CSF) and can be assessed (Dubois et al. 2007).

Although there is a great amount of literature on ex vivo samples of Aβ fibres and its relation to Alzheimers disease, the literature on *in vitro* Aβ fibres is lacking. Nevertheless, it was discovered that Lomakin *et al.* were able to produce Aβ fibres in 1996 using 0.1M HCl. The fibres were produced using $A\beta_{1\text{-}40}$ at room temperature. They were measured using lasers and light scattering techniques alongside CD measurements. Unfortunately, the exact conditions used to create these fibres - pH, agitation etc - are not detailed in this paper making the results difficult to compare with fibre works from other proteins.

Aβ Preparation; Recombinant Protein Expression

Unlike insulin, Aβ is an expensive protein to obtain from manufactures. The price of 50mg of Insulin is equivalent to only 1mg of Aβ thus not providing

much freedom for experimental trial and error. Due to this, it was decided to express and purify Aβ ‘in house’.

Considering the size and complexity of protein molecules, chemical synthesis is not a practical approach for ‘in house’ protein manufacturing. Instead, molecular biologists have developed a way in which to prompt living cells into producing functional proteins from genetic (DNA) templates. This process is referred to as recombinant protein expression.

Recombinant protein expression is the process of synthesising large amounts of a specific protein from DNA into which the gene sequence that codes for the protein has been artificially introduced into a host organism. DNA, unlike protein, is relatively simple to construct *in vitro* as it is only composed of 4 nucleotides instead of 20 amino acids. The development of recombinant protein techniques has revolutionised biochemistry. Before, huge quantities of animal and plant tissue were needed to purify small amounts of protein. This was not only wasteful and a time consuming process, but also did not always guarantee high yields of protein sample. The ability to express and purify proteins in large quantities allows for its biochemical characterization, its use in industrial processes and the development of commercial goods.

At the theoretical level, the steps needed for obtaining a recombinant protein are quite straightforward. You take your gene of interest, clone it in an expression vector, transform it into the host of choice, then induce, purify and characterise. However, in practice, things can often go wrong. It is relatively common to witness poor growth of the host, the formation of inclusion bodies, protein inactivity, and even not obtaining any protein at all (Rosano and Ceccarelli 2014). Choosing the host cell in which protein production will take place is the most important step as it will initiate the outline of the process and define the technology needed.

The process of protein expression can be accomplished using either prokaryotic or eukaryotic cells including bacteria, yeast, filamentous fungi, and unicellular algae. Each of these cells has its own set of advantages and

disadvantages. For example, prokaryotic systems such as bacteria are easy to culture, quick to grow and produce high yields of recombinant protein. Unfortunately, many eukaryotic proteins produced in prokaryotic cells are inactive as the bacteria are unable to perform post translational modifications required for protein folding. This often results in the protein becoming an insoluble inclusion body which can be difficult to recover without the use of strong denaturants. On the other hand, eukaryotic systems will often produce fully functional proteins however the cells take much longer to grow, the yield is lower and production is costly compared to the prokaryotic counterparts (Andersen and Krummen 2002).

Escherichia coli (*E.coli*) bacterium is one of the organisms of choice for the production of recombinant proteins. Its use as a cell factory is well-established and it has become the most popular expression platform. The advantages of using *E.coli* as the host organism are well known. Transformation with exogenous DNA is fast and easy, plasmid transformation of *E. coli* can be performed in as little as 5 min and rich complex media can be made from readily available and inexpensive components. *E.coli* has unparalleled fast growth kinetics and it achieves high cell density cultures easily. *E.coli* cells have a doubling time of only 20 minutes when in glucose-salts media in optimal environmental conditions meaning that a culture inoculated with a 1/100 dilution of a saturated starter culture may reach stationary phase in only a few hours(Sezonov, Joseleau-Petit, and D'Ari 2007). However, it should be noted that the expression of a recombinant protein may impart a metabolic burden on the microorganism, causing a considerable decrease in generation time (E. Bentley et al. 1990).

The second most important step in recombinant protein expression is choosing your expression vector. An expression vector is a plasmid or virus used to introduce the specific protein gene into the chosen target host cell. Examples of expression vectors include the pET , pUC, pBAD and pQE series. In fact, there is an enormous catalogue of available expression vectors in use today. They are a result of multiple combinations of replicons, promoters, selection markers, multiple cloning sites, and fusion protein/fusion

protein removal strategies. In order to make an informed decision when choosing an appropriate expression vector, one must evaluate each feature according to one's individual needs:

Replicons

Replicons – the first feature – are a region of DNA or RNA that replicate from a single origin of replication. Genetic elements which undergo replication, such as plasmids, always contain a replicon and its associated *cis*-acting control elements. It is essential to have some form of control on plasmid replication as a high plasmid number may impose a metabolic burden that decreases the bacterial growth rate and may produce plasmid instability. This would cause the number of healthy organisms for protein synthesis to fall and would result in a decrease in production yields (E. Bentley et al. 1990). Commonly used vectors such as the pET series possess the pMB1 origin, a derivative of Co1E1 which allows 15-60 copies per cell. Alternatively, the pUC vector series contains a mutated form of the pMB1 origin which allows between 500 and 700 copies per cell (Bolivar et al. 1977) (Minton 1984).

Promoters

Promoters – the second feature - are DNA sequences located in the 5' region adjacent to the transcriptional start site. RNA polymerase and accessory proteins (transcription factors) bind to the promoter to initiate production of an mRNA transcript. The T7 promoter system present in the pET vectors is extremely popular for recombinant protein expression. In this system, the gene of interest is cloned behind a promoter recognized by the phage T7 RNA polymerase. This highly active polymerase should be provided in another plasmid or, most commonly, it is placed in the bacterial genome in a prophage (λDE3) encoding for the T7 RNAP under the transcriptional control of a *lac*UV5 or *lacO* promoter. Thus, the system can be induced by lactose or its non-hydrolyzable analog isopropyl β-D-1-thiogalactopyranoside (IPTG). (Studier and Moffatt 1986). T7 lysozyme binds to T7 RNAP and inhibits transcription initiation from the T7 promoter (Stano and Patel, 2004).

After induction, the amount of T7 RNAP produced surpasses the level of polymerase that T7 lysozyme can inhibit. The “free” T7 RNAP can thus engage in transcription of the recombinant gene (Dubendorff and Studier 1991).

Selection markers

The selection markers are responsible for deterring the growth of plasmid-free cells. In the *E.coli* system, antibiotic resistance genes are habitually used for this purpose. Resistance to ampicillin is provided by the *bla* gene in the plasmid whose product is a periplasmic enzyme that inactivates the β-lactam ring of β-lactam antibiotics. During recombinant protein expression, ampicillin is added to the culture. Those cells which have incorporated the antibiotic resistant plasmid exhibit a continuous secretion of β-lactamase. Over time this degrades the antibiotics and within a couple of hours, ampicillin is almost depleted. Those cells which have not incorporated antibiotic resistant plasmids do not have the ability to inactivate the β-lactam ring of β-lactam antibiotics and are killed (Korpimäki, Kurittu, and Karp 2003).

Affinity Tags

It is crucial when expressing a recombinant protein to have a means of purifying it from the *E. coli* cellular milieu. It can be found that purifying a protein via such characteristics as size and charge, although effective, can often lead to multiple rounds of purification due to the likelihood of another protein in the milieu having similar properties. Each round of purification is a time consuming task and can lead to delays in actual experimentation of the protein. However, more recently scientists have found that fusion of the recombinant protein with a stretch of amino acids can be used to allow single step purification.

The first affinity tags used were large proteins utilized almost exclusively for protein expression and purification in *E.coli*. Examples of these include staphylococcal protein A; a polypeptide 280 residues long which can be

purified using its affinity to IgG Sepharose, and lacZ, a peptide 1024 residues long which can be purified using its affinity to immobilized *p*-amino-phenyl-β-D-thio-galactosidase (APTG). Both of these affinity tags were successful with respect to single step affinity purification, however, their large peptide sequence often caused problems within the process of protein expression itself. Their large sizes influenced the proteolytic stability, increased the solubility and/or expression of the heterologous proteins and occasionally resulted in denaturation and loss of activity of the fusion protein (Kimple, Brill, and Pasker 2013).

Since this time, advances in recovery technique have improved the ability to isolate Protein A and lacZ and their associated protein complexes without any complications in protein expression and activity. This allows for the continued use of these affinity tags today. However, many other affinity tags have recently been developed which improve on the two original designs. These include the polyhistidine (His) tag, maltose binding protein, Calmodulin Binding Peptide and Streptavidin/Biotin-based tags. Each of these is isolated through their affinity for metal, amylose, calmodulom and avidin respectively. They are all small peptide tags which make them less likely to interfere with protein folding however this can still happen on occasion. They can be tagged on either the N- or C- terminal end of the protein although it is preferable to place them on the latter if a signal peptide is present on the N-terminus. Also, if the three dimensional structure of the protein is known, it is wise to place the affinity tag on the end which is solvent accessible instead of the end which is buried so to minimise the affect it has on tertiary structure. Ultimately it is best to remove the tag in enzyme digestion processes following purification so not to interfere with the protein activity and structure (Rosano and Ceccarelli 2014).

Of these more modern affinity tags, the his-tag is the most widely used. His-tags were first used in single step protein purification in 1991 to purify recombinant galactose dehydrogenase by immobilized metal affinity chromatography (Lilius et al. 1991). They are small peptide sequences usually consisting of 6 histidine residues although they can be anything between 2

and 10 residues. His-tags exhibit strong affinity to transition metal ions in particular with cobalt, copper, nickel, zinc, calcium and iron ions and form strong co-ordination bonds with each although nickel is the most commonly used.

Affinity Chromatography

Affinity chromatography is the technique used to separate affinity tagged recombinant proteins from mixtures through specific ligand binding. This is achieved through use of a column containing immobilised versions of the ligand specific to the tag. In the cases where his-tags are used, an affinity column containing nickel is typically chosen. Here, when a mixed solution is passed through the column only those with an attached his-tag will show binding affinity towards the nickel in the column and bind. The remaining protein will simply pass through the column and be eluted. Following the elution of the unwanted material, the bound, his-tagged protein is recovered using imidazole. Imidazole is an organic compound with the formula $C_3N_2H_4$. When solubilised in water imidazole has a higher affinity to nickel than the his-tag and displaces it allowing the elution of the recombinant protein.

Following the purification of a recombinant protein it is always vital to confirm its purity and identity before any subsequent experimentation. The purity of a protein is typically confirmed using something called SDS PAGE (sodium dodecyl sulfate polyacrylamide) electrophoresis and the identity of a protein by a technique called Western Blotting.

Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis (SDS PAGE) Electrophoresis

SDS PAGE electrophoresis is a process which separates protein mixtures via their electrophoretic mobility; a quality that is affected by length, conformation and charge of a protein. In order to eliminate two of those variables, the protein samples are heated with sodium dodecyl sulphate (SDS). SDS is an anionic denaturant which breaks down protein structures into their monomeric state. Heating with SDS linearises proteins and binds to them giving an even distribution of charge per unit mass; *i.e.* the longer the protein, the more SDS that will bind and the higher the negative charge. When carrying out electrophoresis, the protein samples must be pipetted into an acrylamide gel and an electric field must be applied. When this happens the proteins, which are negatively charged due to the SDS, are attracted to the positively charged anode and begin to migrate through the gel. The bigger proteins, although having a greater charge, are unable to travel through the tightly packed matrix of the gel so do not migrate as far as the smaller proteins.

Western Blott

The identity of a recombinant protein is confirmed using Western blotting. Western Blotting is a technique which identifies a specific protein from a mixture through the use of antibodies specific to it (Tahrin Mahmood 2012).

The first stage – separation of the proteins - is carried out through the use of SDS PAGE electrophoresis as explained above. Samples are denatured, coated in negative charge and then subject to an electric current. Smaller proteins, although not as negatively charged, travel faster through the porous gel towards the positively charged anode. Larger proteins, due to their size, are unable to travel efficiently through the gel thus proteins are separated via their molecular weight.

Immediately following protein separation, step two is performed: transference onto a solid membrane. There are a variety of methods that can be used for this process, including diffusion transfer, heat-accelerated convectional transfer and vacuum blotting transfer. However, due to its speed and transfer efficiency, the electrophoretic transfer method is the most commonly used today. Here, the proteins are 'pulled' out of the acrylamide gel onto a membrane using a current that runs perpendicular to the gel. In doing so, the separated proteins are able to maintain their organisation during the transfer. Transfer efficiency varies greatly amongst proteins as each has a unique ability to migrate out of the gel and bind to the membrane. The efficiency of transfer depends on factors such as the gel composition, field strength, transfer time, complete contact of the gel with the membrane and composition of the proteins (Overview of Western Blotting, 2017).

The use of a membrane is essential in order to make the proteins accessible for antibody binding. It can be made from either nitrocellulose or polyvinylidene difluoride (PVDF) as both materials have non-specific binding properties. This characteristic allows all proteins to bind equally well to them through hydrophobic and charged interactions. Unfortunately, the antibodies used in the final step of western blotting – antibody staining – are proteins themselves so would bind non-specifically to the entire surface of the membrane. As a way of preventing this, the membrane is soaked in a dilute solution of protein, usually 3-5% non-fat milk, plus 0.1% of Tween 20 detergent in a process called 'blocking'. Blocking is an inexpensive method of attaching protein to the remaining space on the membrane. Thus, when the antibody is finally added, it will only bind to its specific protein – the target protein (Jensen 2012).

The antibodies used in western blotting are modified. They are linked to a reporter enzyme such as alkaline phosphatase or horseradish peroxidase which, when exposed to an appropriate substrate, produces a colour. At present both specificity and colour can be achieved using a single antibody, however, these roles were originally split between two: primary and secondary antibodies. The primary antibodies were specific to the target protein making

theirs a specific choice for the target protein. In general, primary antibodies were not directly detectable so tagged secondary antibodies would have to bind to the primary for detection purposes. Antibodies are able to bind to another of its kind *i.e.* bovine originated antibodies bind to other bovine antibodies. Therefore, in this case, the secondary antibody would have to be anti-bovine. This 'same-species' quality means that the same secondary antibodies can be used for various western blotts regardless of the protein involved. Although successful, the two step staining method is susceptible to non-specific binding and involves additional incubation steps compared to the single step method. Having both the specificity and the reporter enzyme on the same antibody, although more costly, reduces experimental time and reduces the amount of reagents necessary (Jensen 2012) (Tahrin Mahmood 2012).

Visualisation of the target protein occurs upon the incubation of the membrane with a substrate that reacts with the reporter enzyme *e.g.* peroxidase. This converts the soluble dye into an insoluble form of a different colour. It precipitates next to the enzyme and in doing so stains the membrane. The membrane can then be compared to the original SDS PAGE gel and the identity of the target protein can be calculated through its molecular weight. Alternatively, fluorescently tagged antibodies can be used as well as chemiluminescent substrates. The resulting light output can be measured by a fluorescence imaging system or a film/ charge-coupled device (CCD) camera respectively.

Methods

Aβ Expression

The gene for Aβ42 was cloned and inserted into a pET15b vector by a predecessor of the 'Crystallography' lab at the Royal Free Hospital, Hampstead Heath. The pET15b vector sequence is 5708 base pairs long and contains ampicillin resistance. Additional features of this vector include a T7 promoter at residues 463-479, T7 transcription start site at residue 452, His-Tag coding sequence (362 – 380), Multiple cloning sites (Nde1-BamH1) (319-335), T7 terminator (213-259), Lac1 coding sequence (866-1945), pBR322origin (3882) and a bla coding sequence at residues 4643-5500 **(Figure 47)**.

A 1µl measurement of $A\beta_{42}$ genes was used to transform 40µL of BL21 (DE3) competent *E.coli* cells containing chloramphenicol resistance. This was achieved through incubation of the genes and cells for half an hour on ice. They were then heat shocked for 45 seconds at 42˚C and later recovered on ice for a further 10 seconds. Following transformation the cells were plated onto luria broth (LB) agar plates containing ampicillin and chloramphenicol antibiotics and incubated overnight at 37˚C.

Successful colony growth occurred in the *E.coli* cells which had incorporated a vector and thus contained resistance to both antibiotics. A colony was picked at random and transferred to 50ml of antibody containing LB for overnight incubation with 120rpm shaking at 37˚C. The following day, 1ml of the new bacterial growth was pipetted into 1 litre of autoclaved Overnight express Medium (contents on 'Buffer Solutions' page) and incubated overnight a final time at 37˚C whilst shaking at 120rpm.

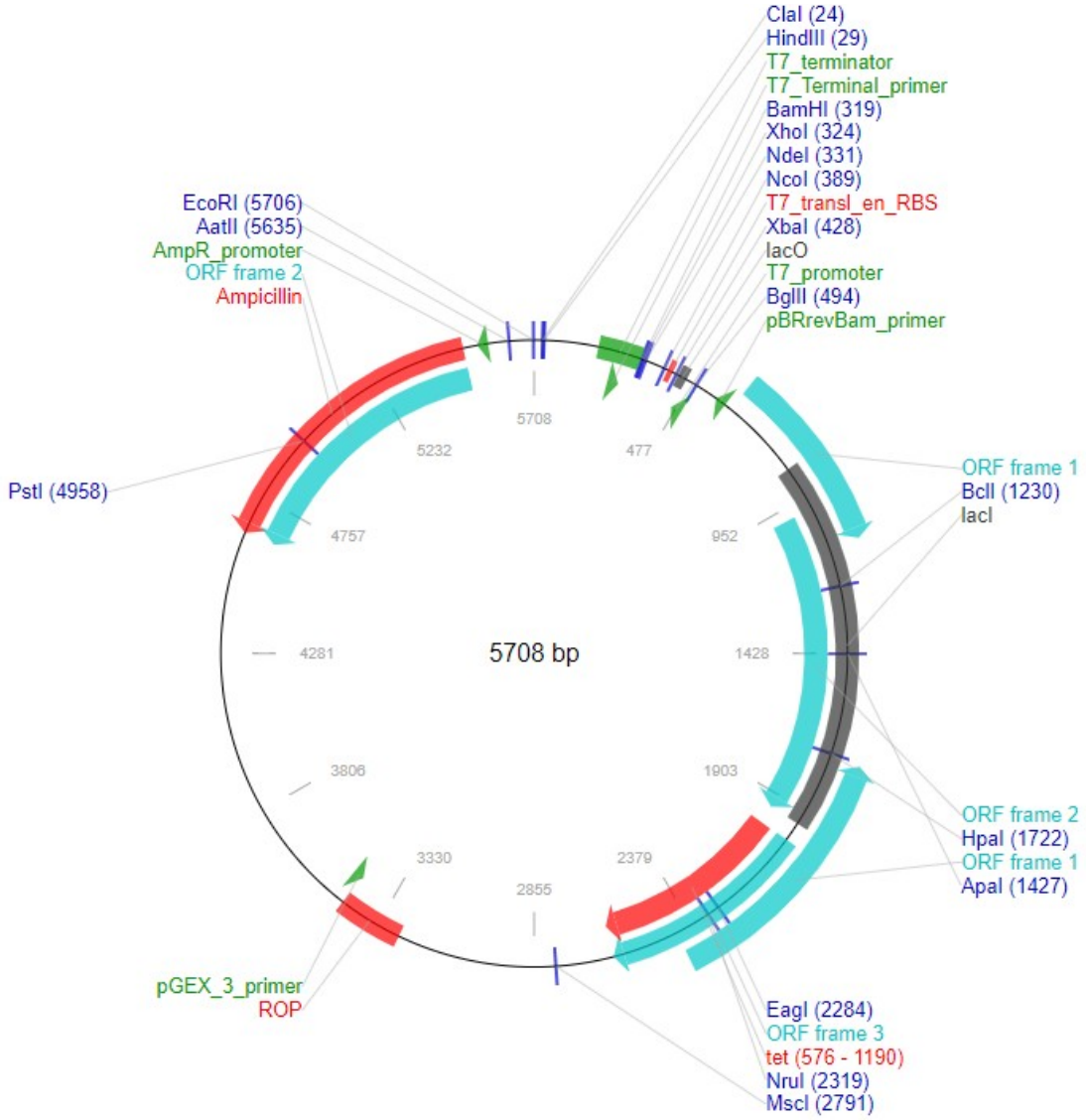


Figure 47: Simplified diagram of a pET15b vector. The vector sequence carries an N-terminal His-Tag followed by a thrombin site and three cloning sites. pET15b is 5708 base pairs long and contains ampicillin resistance. It has a 5' sequencing 1 primer of T7Fwd of sequence 5'd[TAATACGACTCACTATAGGG]3'. The pET15b vector landmarks are as follows:

T7 promoter	***463-479***
T7 transcription start	***452***
His-Tag coding sequence	***362 - 380***
Multiple cloning sites (Nde1-BamH1)	***319-335***
T7 terminator	***213-259***
Lac1 coding sequence	***866-1945***
pBR322origin	***3882***
bla coding sequence	***4643-5500***

Aβ Purification

Upon completion of bacterial growth the cells were lysed on ice with sonication. Four rounds of sonication were carried out on a Branson Sonifier Cell Disruptor 2000. Each round included three, 75 seconds pulses with 30 second pauses in between. Between each round the sample was subject to spinning via centrifugation at 15,000 rpm, 4˚C for 10 minutes. Following each spin, the supernatant was discarded so to remove unwanted cellular components such as the nucleus, mitochondria and other cell waste, whilst the remaining pellet was re-suspended in purification buffer. Following the final spin the pellet was re-suspended in 20ml 8M urea so to solubilise the $A\beta_{42}$ inclusion bodies. Once solubilised, the sample was diluted to 2M urea at which point the sample was ready for purification via affinity chromatography.

A 1ml HiTrap nickel column was used for recombinant $A\beta_{42}$ purification. Initially, the column was equilibrated using 4ml of sample buffer. Following this, the mixed sample was passed through twice to ensure his-tag-nickel-ion binding and the remaining, unwanted, protein was eluted and discarded. The application of 4ml low strength imidazole (50mM) removed any low-specific nickel binding protein and 300mM imidazole was later used to detach the remaining Aβ from the column.

The purity of the product was tested on a pre-manufactured Novex™ WedgeWell™ 4-20% Tris-Glycine gradient gel run in 1x running buffer at 25-30mA for up to 2 hours. Once removed from the casket, the gel was stained for 40 minutes in coomassie and left to de-stain overnight in de-stain buffer. (Buffer contents on ‘Buffer Solutions’ page)

Western Blotting

For the purpose of the western blott it was decided to use 16% acrylamide gels instead of the normal 12% due to the small size of the protein involved. On the original 12% gels it was noted that the smaller molecules would “run off” the gel if

left under current for too long. Contrastingly the bands would be difficult to distinguish if not left for long enough. By using a gel with a larger percentage of acrylamide, the proteins were presented with a denser environment, slowing their progression and allowing a more defined gel. A common problem with these denser gels, however, was the increased running time and increased risk of damaging the gel from the necessity of higher currents.

The contents of the 'purified Aβ' sample were separated through gel electrophoresis simultaneously on two separate gels. One of these was used for membrane transference whilst the other was coomassie stained so to be used as a visual comparison later. The gel which was transferred to the nitrocellulose membrane was done so immediately after electrophoresis through the use of a perpendicular current. The electric field caused the negatively charged proteins to move towards the anode onto the membrane whilst maintaining the organisation they had on the gel. The transfer was run at 100V for 1.5-2hours on ice creating a mirror image of my gel.

Once transferred, the membrane was blocked with 5% milk in Tris Buffered Saline Tween (TBS-Tween). Blocking was performed for one hour. Following blocking, an antibody specific to His-tags was added. The antibody also contained the enzyme horseradish peroxidase (HRP) which, when stained with acetate buffer, AEC chromogen and hydrogen peroxide at a ratio of 2:1:1 respectively, became a red colour. Following an hour long antibody incubation, all solutions were washed off and the membrane was stained with the above mixture.

Aβ Preparation

Aβ elutions from each expression attempt (table?) were combined and concentrated to make a stock of approximately 1.5ml of 3.41mg/ml solution. This stock was dialysed overnight against neutral "TE buffer" (20mM TRIS, 150 NaCl, 0.01% Na Azide, pH8) at a ratio of 1:2000 sample to dialysing buffer respectively, in order to remove the urea from the solution. The presence of urea in the

fibrillisation assays containing SAP would have had an effect on the structure of the SAP and thus altered the results. The urea was no longer necessary as the oligomers had been solubilised and removal of the denaturant should allow the Aβ to revert back to its natively unfolded form.

Aβ Fibrillogenesis

The rotator method as described in chapter 2 was carried out using the expressed $A\beta_{42}$. The protocol was carried out exactly with the exception of a lower protein concentration (0.3mg/ml) due to limited stocks of $A\beta_{42}$. Having already performed fibrillogenesis with insulin at lower concentrations it was known that fibre mass is still able to form in these conditions and that its ThT fluorescence is proportional to the amount of protein initially added.

In order to set up the fibrillisation assays, the Aβ stock was diluted to a 1ml sample of 0.3mg/ml. It was chosen to continue using the 1ml sample size standard to maintain the same surface area during tumbling and thus not introduce a new variable to the experiment which could affect results. This was achieved using TE buffer in a 1.5ml eppendorf. The samples were continuously inverted – or “flipped” - at a speed of 20rpm on a revolving Stuart Rotator SB3 for 10 days at room temperature as per the protocol fro making Insulin fibres described in chapter 2. This time, a variety of concentrations was not set up due to the limited stock amounts. It had already been seen that at least, in the case of insulin, that fibre masses were proportional to protein concentration. That assumption was applied to these assays also.

Fibre Formation with SAP decamers

In order to observe the effect of SAP decamers on Aβ fibre growth, the rotator method was repeated with the addition of 0.06mg/ml SAP. This maintained the stoichiometric ratio 1:100 of SAP decamers to Aβ respectively, as carried out with the insulin fibres. Each sample was continuously inverted at a speed of 20rpm on a revolving Stuart Rotator SB3 for 10 days at room temperature. This protocol was repeated for both the L55P and S52P variants.

Sample	Aβ (3.41 mg/ml)	TE Buffer	SAP (6.66mg/ml)	dH_2O	Final Aβ Concentration
1a	88 µl	903 µl	-	9 µl	0.3 mg/ml
1b	88 µl	903 µl	9 µl	-	0.3 mg/ml

Fibre Formation with SAP Pentamers

To perform this particular assay the SAP was diluted to 0.06mg/ml using TC buffer. This was left to incubate for half an hour to ensure saturation of SAP calcium pockets. Following this, 0.3mg of Aβ was added and the samples were continuously inverted at a speed of 20rpm on a revolving Stuart Rotator SB3 for 10 days at room temperature.

Sample	Aβ (3.41 mg/ml)	TC Buffer	SAP (6.66mg/ml)	dH_2O	Final Aβ Concentration
2a	88 µl	903 µl	-	9 µl	0.3 mg/ml
2b	88 µl	903 µl	9 µl	-	0.3 mg/ml

Control

To see the effect the salt was having on Aβ fibre formation the assay had to be repeated in a High Salt buffer (“THS Buffer” - 20mM TRIS, 600mM NaCl, 0.01% Na Azide, pH8). A concentration of 0.3mg/ml Aβ was continuously inverted at a

speed of 20rpm on a revolving Stuart Rotator SB3 for 10 days at room temperature in the presence and absence of 0.06mg SAP.

Sample	Aβ (3.41 mg/ml)	THS Buffer	SAP (6.66mg/ml)	dH_2O	Final Aβ Concentration
3a	88 µl	903 µl	-	9 µl	0.3 mg/ml
3b	88 µl	903 µl	9 µl	-	0.3 mg/ml

Aβ Fibre Analysis using ThT

Please refer to the 'Fibre Analysis using ThT' section in chapter 2.

Fibre Detection using Congo red

Please refer to the 'Fibre Detection using Congo red' section in chapter 2.

Fibre Analysis using Circular Dichroism

Please refer to the 'Fibre Analysis using Circular Dichroism' section in chapter 2.

Results and Discussion

Protein Expression and Purification

The $A\beta_{42}$ clone was expressed in BL21 (DE3) competent *E.coli* cells then cultured and purified as per the methods described above. Throughout the course of protein expression the progression of the purification process and the purity of the final product was analysed via SDS-PAGE electrophoresis on a pre-manufactured Novex™ WedgeWell™ 4-20% Tris-Glycine gradient gel. The contents of each lane is as follows: Lane 1) Pre-stained molecular marker, 2) Sample of lysed cells, 3-6) Sample after first, second, third and forth round of centrifugation and 7) Purified protein following elution from a nickel column **(Figure 48)**.

Upon analysis of the protein expression gel one can see that lanes 2 – 6, which represent the sample during the sonication process, each contain multiple bands. These represent the vast array of bacterial protein content as is expected during this part of the purification process. Lane 7 of this gel represents the sample after submission to affinity chromatography. This lane contains a single, well defined band at approximately 36kDa as well as various very faint bands at approximately 39, 47, 52, 80 and 84 kDa.

The contents of lane 7 draw attention to two matters. The first concerns the molecular weight of the strong band and the second concerns the presence of other bands in that lane.

With regards to the first issue, one can see that the molecular weight of the strong band, 36 kDa, is not representative of $A\beta_{42}$ which has a molecular weight of only 6kDa. This suggests that $A\beta_{42}$ is not present in the sample. Perhaps it was lost or discarded at some point during the purification process or perhaps it was run off the gel due to excessive electrophoresis times. The latter of these seems most likely as molecular weights lower that 20kDa cannot be seen.

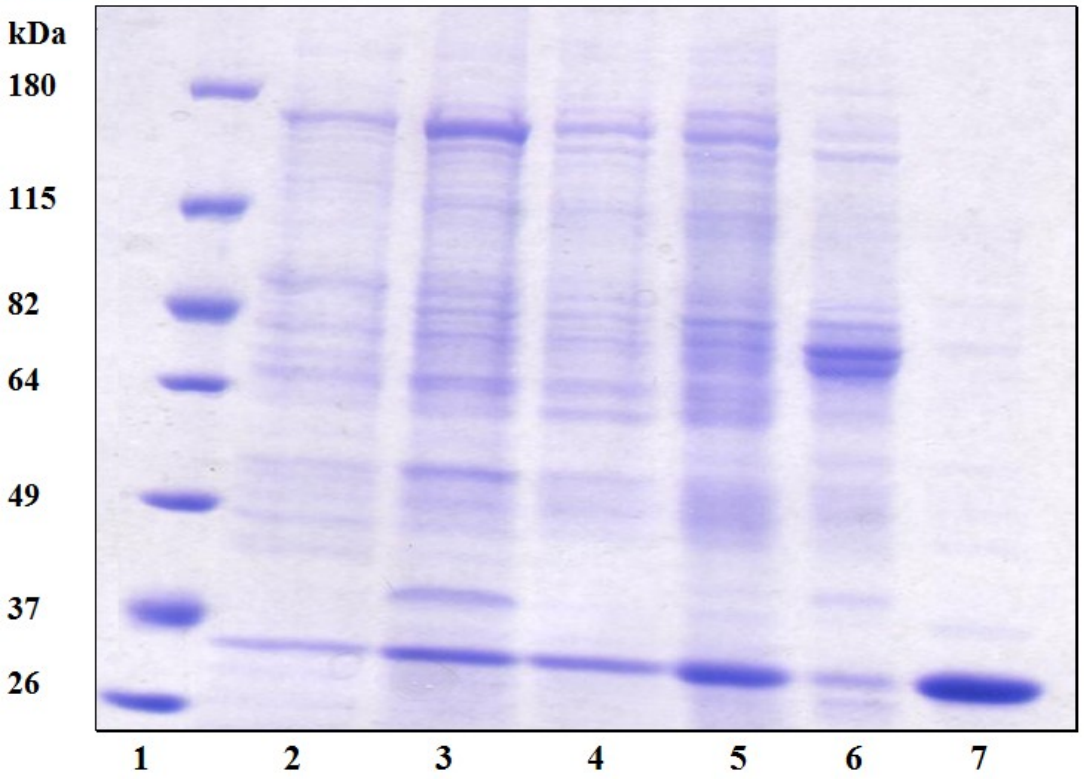


Figure 48: Electrophoresis gel showing the progression of Aβ purification in a Novex™ WedgeWell™ 4-20% Tris-Glycine gradient gel. A total of 10µl was added to each lane except in the case of the marker where only 5µl was added for preservation purposes. During preparation, samples were diluted to 1mg/ml using laemmli buffer containing sodium dodecyl sulfate (SDS) and boiled in order to enforce monomerisation and uptake of negative charge. Electrophoresis was carried out for 1 hour 20 minutes at 25 mA. The contents of each well are as follows: 1) Pre-stained molecular marker. 2) Sample of lysed cells. 3) Sample after first round of centrifugation. 4) Sample after second round of centrifugation. 5) Sample after third round of centrifugation. 6) Sample after last round of centrifugation. 7) Purified protein following elution from a nickel column.

With regards to the second issue, the presence of multiple bands in a purified lane indicates that the protein is not pure in content. The presence of additional bands could be a result of sample contamination from either human error or non-specific nickel binding. It is entirely possible that the nickel column had lost its specificity. This is a common problem following multiple uses of the same column. Another explanation could that the $A\beta_{42}$ had indeed not been expressed and that the bands in the gel represented other proteins with mild affinity for nickel binding.

Ideally, in order to test all these theories, the entire protein expression protocol would need to be repeated. However, this is a very long and wasteful process so instead it was decided to use a much quicker method first, SDS-PAGE electrophoresis. By producing another gel, one run for a shorter period of time, one would get an indication as to whether any $A\beta_{42}$ is present in the sample at all. The brightness of the band would then provide an indication as to whether affinity binding had occurred or not. As a way to resolve the apparent impurity of the sample, a round of affinity chromatography would need to be carried out first with a new nickel column.

Upon producing this gel, however, the exact same results were observed. A strong band was present at approximately 36kDa alongside various very faint bands at approximately 39, 47, 52, 80 and 84 kDa. This lead to the following conclusions: Plasmid transformation had indeed been successful. The antibiotic resistance exhibited by the bacterial cells proves this. His-tag affinity had also been successful. The presence of a strong band in lane 7 demonstrates this. What is not known is the identity of the protein that had been expressed.

Considering that the vector used during this protocol was created by an unknown predecessor, it as decided to submit the clone through DNA sequencing in order to verify its contents. The result of the DNA sequencing of the $A\beta_{42}$ vector was translated via the translation tool available on the ExPasy website and compared with the $A\beta_{42}$ sequence available on the ncbi website (PDB code 2MXU_L). The 5'3' open reading frame 2 illustrated that $A\beta_{42}$ protein was indeed present and that it did contain the His-tag required for nickel binding and single step

purification. This indicated that the SDS PAGE gel anomalies were a result of protocol negligence alone **(Figure 49)**.

Further reading of the literature, however, led to the discovery that Aβ is prone to aggregation and oligomer formation when expressed in bacterial cells as experienced by Lesné *et al*. in 2006. The oligomers are complexes made from multiple Aβ monomer units and are usually exhibited as structures made up of multiples of three. They cannot be separated by either urea or SDS meaning they would appear on a SDS-PAGE gel as multiple sized bands instead of as a single monomerised unit. For example, a single Aβ monomer has the molecular weight of approximately 6kDa so oligomers composed of 3, 6, 9 etc monomers would exhibit weights of approximately 18, 36 and 54kDa respectively on an SDS electrophoresis gel. This is in agreement with Figure 48. The strong band at 36kDa would indicate that the oligomer approximately 36kDa in size - the one containing 6 Aβ monomers - is the most prominent form.

When looking at the fainter bands in this gel one will notice that the molecular masses of these bands do not necessarily add up precisely to units of 15kDa, this can be explained by the fact that the oligomers are not linear; they are globular. Globular proteins experience a retarded migration through acrylamide gels due to their bulk in size. The larger the protein, the more difficulty it experiences when travelling through the gel matrix. The protein marker used in lane 1 consists of linear proteins. These do not experience the same restrictions upon travelling down an acrylamide gel as globular proteins so therefore travel at a faster rate. This means that the size markers demonstrated at particular distances in lane 1 are not necessarily representative of the molecular weights of the globular proteins present in lane 7.

This oligomerisation phenomenon could be the result of a number of reasons. The introduction of a foreign gene into *E.coli*, can cause the spatio-temporal control of its expression to be lost. The newly synthesized recombinant polypeptide is being expressed in the microenvironment of *E. coli*, which may differ from that of the original source in terms of pH, osmolarity, redox potential, cofactors, and folding mechanisms thus encouraging aggregation of some sort.

One must also take into account that in high level expression, hydrophobic stretches in the polypeptide are present at high concentrations and available for interaction with similar regions. Prokaryotic chaperones do not always have the capability to fold eukaryotic proteins into their native state. Even if they do, they are not present at a high enough concentration to fold this excess number of recombinant protein molecules. All of these factors lead to protein instability and aggregation (Hartley and Kane 1988).

Verification of the presence of Aβ in the 'purified' sample was carried out using Western blotting in conjunction with his-tag specific antibodies. This was done alongside proteins lacking any his-tags so to act as a control. The contents of each gel are a follows: 1) Purified Aβ with his-tag attached. 2) Albumin. 3) TTR. 4) SDS PAGE/Western Blott Marker.

When looking at the results of the Western blott various things can be observed. Firstly, the high acrylamide percentage prevents the protein samples from travelling through the gel in a neat and clear manner. This is true even after being submit to a current for two hours at a time. Unfortunately this could not be helped due to the small size of the protein in question. A higher percentage gel had to be used to ensure that there was no chance the $A\beta_{42}$ would run off the gel. Although the gel could have been left to run longer in an attempt to separate the bands, by doing so there was an increased risk of damaging the gel **(Figure 50)**.

When comparing the Western blott with its SDS-PAGE counterpart one can see that the contents of lane 1 – purified Aβ - resulted in positive identification of his-tags where as lanes 2 and 3 – albumin and TTR respectively - did not. This proves that the recombinant Aβ protein had successfully been integrated, expressed and purified from *E.coli* cells. The crucial observation is that the western blott caused red bands to appear not on a single band but on each of the anomalous bands seen in the sample. This confirms the theory that the Aβ was indeed pure and simply forming oligomers.

T F P P S **Stop** K **Stop** F C L T L R R R Y T **Met** G S S **H H H**
H H H S S G L V P R G S H **Met** D A E F R H D S G Y E V H
H Q K L V F F A E D V G S N K G A I I G L **Met** V G G V V
I A **Stop Stop** G S G C **Stop** Q S P K G S **Stop** V G C C H R
Stop A I T S I T P W G L **Stop** T G L E G F F A E R R N Y I R
I S R K R P G S T G I T K P **Met** P T A S R V T V P R **Met** T
Met S A L L D F I H G A **Stop** L R **Stop** Q F N C D K L P H
Stop S L S **Met** I S C Q T **Stop** E F L K T K G P R D T P I F I
G **Stop** C H D N N G F L D V R W H F S G K C A R N P Y L
F I F L N T F K Y V S A H E T I T L I N A S I I L K K E E Y
E Y S T F P C R P Y S L F C G I L P S C F C S P R N A G E S
K R C **Stop** R S V G C T S G L H R T G S Q Q R **Stop** D P
Stop E F S P R R T F S N D E H F **Stop** S S A **Met** W A R Y Y
P R V D A R A R A T R A A A Y T I L Q

Figure 49: Translated sequence of the pET15b vector. Start codons leading to possible proteins are represented in grey and the Aβ sequence is highlighted in blue. The Aβ sequence has a molecular weight of 6.81kDa (including the methionine and his-tag) as calculated using the 'Compute pI/ Mw' programme on the ExPASy website (https://www.expasy.org/).

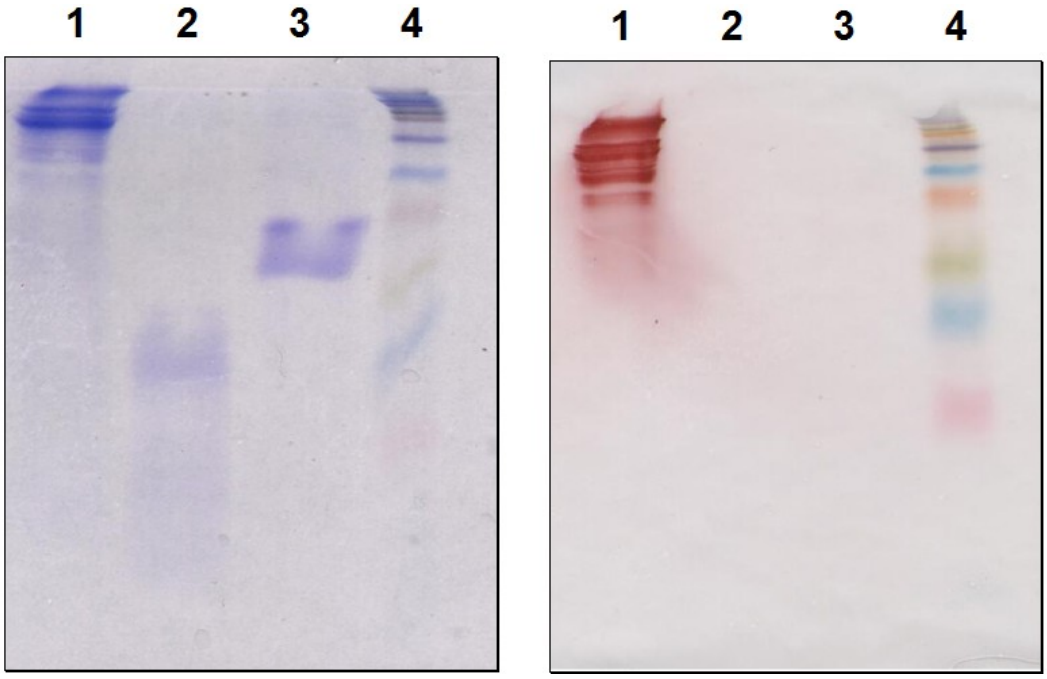


Figure 50: Comparative electrophoresis (left) and western blott (right) results. The contents of each lane are as follows: 1) Purified Aβ with his-tag attached. 2) Albumin. 3) TTR. 4) Western Blott Marker. Proteins were run on home made 16% acrylamide gels at 25mV for 1 hour 45 minutes. The transfer to the nitrocellulose membrane was run at 100V for 1.5 hours on ice. Once transferred, the membrane was blocked with 5% milk and an anti-his tag antibody containing the enzyme horseradish peroxidase (HRP) was used.

Following this revelation, multiple rounds of expression and purification were carried out. In order to encourage monomer formation, the purified solutions were re-submit to 8M urea and shaken overnight at 50rpm. The finishing products of each attempt were combined to produce 1.2ml of stock solution measuring a concentration of 5.2mg/ml.

Aβ Recombination Number	Final Aβ volume (µl)	Final Aβ concentration (mg/ml)	Total Aβ content (mg)
1	200	2.22	0.44
2	200	0.50	0.10
3	200	2.44	0.49
4	200	3.41	0.68
5	200	3.36	0.67
6	100	9.76	0.97
7	100	10.23	1.02

A table of Aβ recombination attempts. Aβ concentrations were determined through use of a nano drop spectrophotometer set at a wavelength of 280nm. Units were converted using the Beer Lambert Law. The finishing products of each attempt were combined to produce 1.2ml of stock solution measuring a concentration of 5.2mg/ml.

Aβ Fibrillogenesis

Again, in order to allow for fair comparison, the rotator method described in Chapter 2 was used as a way in which to encourage Aβ fibre formation. The protocol was carried out exactly but with the exception of a lower protein concentration (0.3mg/ml). This was due to the limited supply of Aβ stocks. Having already performed fibrillogenesis with insulin at lower concentrations it was known that fibre mass is still able to form in these conditions and that fibre mass and its corresponding ThT fluorescence is proportional to the amount of protein initially added.

Attempting to create fibres from the Aβ stock proved to be much more difficult than creating the stock itself. Unlike the previous two proteins, the rotator method did not cause precipitation to occur at in the Aβ assays. Neither cross β-sheets or amorphous aggregate were produced. When tested with ThT fluorescence a flat line was produced directly on the base line.

This difficulty in obtaining Aβ fibres is believed to be a result of Aβs tendency to form soluble oligomers in environments lacking high concentrations of denaturant (*i.e* urea). Lashuel *et al.* have already demonstrated that soluble oligomers can affect fibre formation in TTR. It is logical that the same theory can be used for all proteins. The fact that Aβ did not even form a precipitate is likely due to the fact that these assays were done at a lower concentration.

Although it would be possible to make Aβ fibres in conditions of high denaturant or strong, the results would be rendered useless with regards to these experiments. The original reason behind making fibres was to see the affect of SAP pentamers and decamers on their growth. By creating fibres in conditions of high denaturants, it would become impossible to observe this relationship between Aβ and SAP as the denaturant would have an affect on SAP itself. Even proteins with properties of high stability such as SAP will experience some form of structural change in those extreme conditions and thus prevent the pentamers and decamers from carrying out their roles efficiently. This is turn would disallow any comparison with the insulin results obtained in chapter 2. A similar statement could also be said of any attempt to make Aβ in acidic conditions. Although success would more likely be obtained, the acidic conditions would also have an affect on the interaction of SAP with fibres due to SAP denaturation. Not only is this non-representative of inducible conditions of the body but it would also make any result obtained non-comparable to those from the insulin experiments which were done at neutral pH.

Conclusion

The tendency of Aβ to form either soluble or insoluble oligomers is a very common problem experienced in Aβ expression. In fact, Prokaryotic expression and purification of any highly amyloidogenic peptides such as Aβ have all proven difficult due to their small size, their tendency to aggregate and the toxicity of the formed aggregates. The repeated difficulty in obtaining functional Aβ protein explains the high cost when purchasing this product form manufacturers. Phenomenal efforts have been carried out in an attempt to make the recombinant proteins more competent and fit for use. Of these attempts, the use of protein fusions has been the most successful.

Fusions protect the Aβ from proteolysis and enhance its solubility (Sørensen and Mortensen 2005). The expression of $A\beta_{42}$ fused to segments of a surface protein from the malaria parasite *Plasmodium falciparum* (Döbeli et al. 1995), maltose binding protein (Hortschansky et al. 2005), ubiquitin (Lee et al. 2005), GroES-ubiquitin (Shahnawaz, Thapa, and Park 2007), trigger factor-ubiquitin (Thapa et al. 2008), and hen egg white lysozyme (Nagata-Uchiyama et al. 2007) have all been reported. Although each of these has proven successful in obtaining a more solubilised version of the protein, in order to then remove this undesirable tag, specific proteolysis is an inevitable additional purification step in all of these cases. The proteolytic cleavage reaction is cost-intensive, requires time-consuming optimisation and necessitates post-reaction clean-up, which further reduces the attainable yield.

Although a lack of experimental success was experienced with two of the three proteins, I believe enough progress was made with insulin to provide grounds for further testing. The overall aims of the thesis so far were as follows:

- To clarify whether cross β-sheet fibres can be made *in vitro* at neutral pH as so to replicate physiological conditions.

- To establish whether, in these physiological conditions, SAP has to ability to exhibit *both* pro-fibrillogenic and anti-fibrillogenic activities as Hamazaki and Janciauskiene debated.

- If so, to determine whether these two roles are each specific to a particular face on SAP or to a particular configuration.

- Suggest whether therapeutic strategies using the SAP decamer would be against the fibres would be possible.

Thus far these experiments have proven that it is indeed possible to make cross β-sheet fibres at neutral pH. This was shown very clearly using insulin and not so clearly with TTR or Aβ. What TTR *has* shown is that the capability to form fibres is present, but that the optimum conditions to produce a high yield have not been discovered. The TTR assays also highlighted the importance of analysing literature figures with caution so to avoid making incorrect conclusions regarding fluorescence form differently calibrated machinery.

These experiments have also proved that, at least in the case of insulin, SAP does indeed have the capability of exhibiting both pro-fibrillogenic and anti-fibrillogenic activities and not just one or the other as argued by Hamazaki and Janciauskiene. These activities are dependent on the conformation exhibited by SAP at the time. When in its decameric form, SAP acts as a fibrillogenic stabiliser and enhancer. In its decameric form, however, it behaves in a manner similar to a chaperone.

Unfortunately, this theory cannot be verified based on the results from a single protein alone. Other amyloid proteins will need to be assessed. It is possible that Aβ and TTR were not the best choices for SAP analysis and alternatives should be used instead. Aβ is infamous not only for its cost but for its difficulty in being produced using recombinant technology. It often forms oligomers and insoluble aggregates which interfere with experimental techniques. TTR also has the propensity to form soluble aggregates. In addition to this, TTR is a

highly stable protein known to resist unfolding when the wild type is in its native state. Although amyloid prone variants can be used as a way to combat this, they may not necessarily be as adept to forming amyloid prone intermediates as other, less stable, proteins.

Perhaps *in vitro* fibre formation is not as easily obtained as it would appear especially when attempted at near neutral pH. It may be possible that in the case of some proteins, other components are required in order to allow the cross β-sheet to form, without which, fibre assembly is restricted. The lack of extracellular matrix proteins and proteoglycans could hinder a proteins ability to form cross β-sheet fibres thus resulting in the amorphous aggregate that was seen in the TTR cases. Another possibility is that not all proteins have the potential to form amyloid due to their size or chemical composition. Proteins of only 30-40 residues may not be large enough to form the cross β-sheet structure and thus explain the behaviour of Aβ.

Chapter 4

X-Ray Crystallography

X-Ray Crystallography

An Introduction to SAP Structural analysis

Having performed various experiments with respect to the behaviour of SAP, it became prudent to perform structural analysis on the molecule as well. SAP is a critical component of all amyloid deposits *in vivo* and a key factor in preventing its removal from the body. By understanding the mechanism through which SAP interacts with fibres or other key proteins in the serum, one could possibly devise a way in which to control or reduce fibre stability and growth.

As previously discussed, SAP exhibits an anti-opsinin role, *i.e.* it blocks the immune system from interacting with amyloid deposits and prevents an immune response against it. Recent developments by GlaxoSmithKline have led to the production of antibodies which are specific to SAP. These anti-SAP antibodies have the ability to detach SAP from amyloid deposits and thus leave the deposits exposed to the host's immune system. The exact mechanism behind this is not fully understood however clinical trials have shown that following the removal of the deposit bound SAP from the body, deposits are dissociated and fibre mass is reduced.

Although these results appear promising, more work needs to be carried out in order to guarantee safety and effectiveness of the treatment. These trials have shown to decrease fibre mass but they have yet to be successful in complete amyloid eradication. This means that surgery is still an eventuality as there is the potential for fibre re-assembly, redevelopment of symptoms and further progression of the disease.

The first step in trying to improve this situation is by solving the structure of the SAP:antibody complex. The most efficient way in which to do this is through use of x-ray crystallographic techniques. This is a technique which

uses x-rays as a way in which to determine atomic structure of a protein. To this date, a structure of the SAP:antibody complex has not been published. It is not known how or where they bind. This limits the understanding scientists have on the interactions between the two molecules and how detachment from *in vivo* fibres actually occurs. By solving the structure of this complex, the mechanism behind SAP:antibody binding could be distinguished and thus methods could be developed to optimise this interaction. This would allow for more efficient removal of SAP from the deposits, lead to the optimisation of the treatment and thus allow for higher success rates and lower fatality numbers to be obtained.

The Concept of X-Ray Crystallography

For the purpose of chapter 4 and 5 all structural analysis was performed using x-ray crystallography. X-ray crystallography is a valuable experimental technique which can be used to determine the three-dimensional structure of proteins. It provides the atomic detail necessary for understanding protein-ligand interactions, allosteric mechanisms and facilitating drug design.

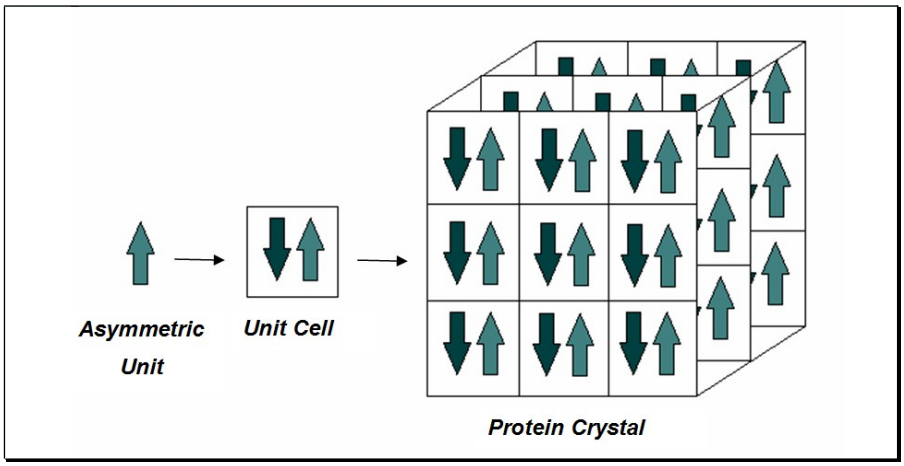
In order to perform x-ray crystallography one must place a pure protein sample in the path of an x-ray beam. Upon doing so, a diffraction pattern is produced in the form of dark spots **(Figure 51, Figure 52)**. The dispersion and intensity of each spot is measured by a detector, transmitted to a computer, and through this, a graphical three-dimensional image of the molecule is reconstructed (Wlodawer, Minor, Dauter, & Jaskolski, 2008). X-ray beams are specifically chosen for this level of analysis as their wavelengths are close in size to the atoms in a protein. The similarity in size guarantees that the two will interact and cause diffraction.

Unfortunately, diffraction from a single molecule does not provide sufficient information for measurement thus there is a need for amplification. This is achieved through conversion of the pure protein sample into a crystal.

Crystals form when the smallest asymmetric sections of a crystal (unit cells) spontaneously assemble into a crystal nucleus. Following this, more unit cells attach to the nucleus to generate an ordered crystalline lattice. Each unit cell packs in identical orientation – similar to the packing of identical boxes in a warehouse – which is essential to prevent canceling out of waves upon diffraction.

X-ray crystallography has grown dramatically in the last 40 years. In 1975 only 78 crystal structures had been published in the Protein Data Band (PDB) whereas now there are more than 100,000. More structures are being published every week than there were per year in 1975. This exponential rate in publishing is due to the advances in computer technologies, the development of more intense x-rays sources and the invention of recombinant molecular biology (Hendrickson, 2000).

Although a useful technique, x-ray crystallography has various challenges which need to be overcome before structure determination can occur. Firstly, the desired protein must be successfully expressed and purified. Once achieved, protein crystals must be grown to an adequate quality and size. This is not always possible or necessarily representative of the protein's natural state. Finally, diffraction must occur at a high enough resolution in order to reveal atomic detail. Each of these steps is incredibly time consuming and can require great amounts of trial and error (G Wagner, S G Hyberts, & Havel, 1992).



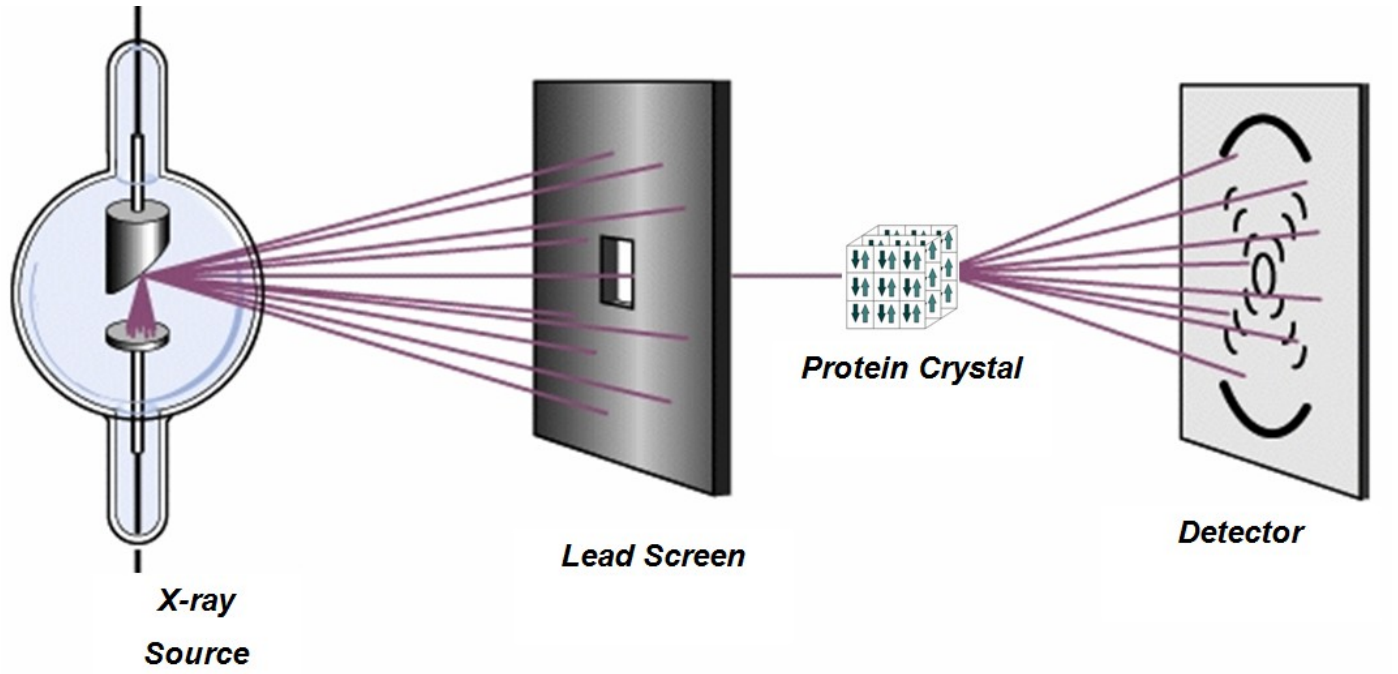


Figure 51: Above – A cartoon representation of a crystal lattice and its constituents (arrows represent protein). Note that a unit cell is the smallest asymmetric section of a crystal. The lattice is formed when identical unit cells assemble spontaneously. Below - A cartoon diagram to demonstrate x-ray diffraction of a protein crystal lattice. An x-ray wave is generated using one of 3 methods: x-ray tubes, rotating anodes or particle storage rings. The generated wave is focused using a lead screen and upon contact with the protein crystal, diffract, generating a pattern of spots.

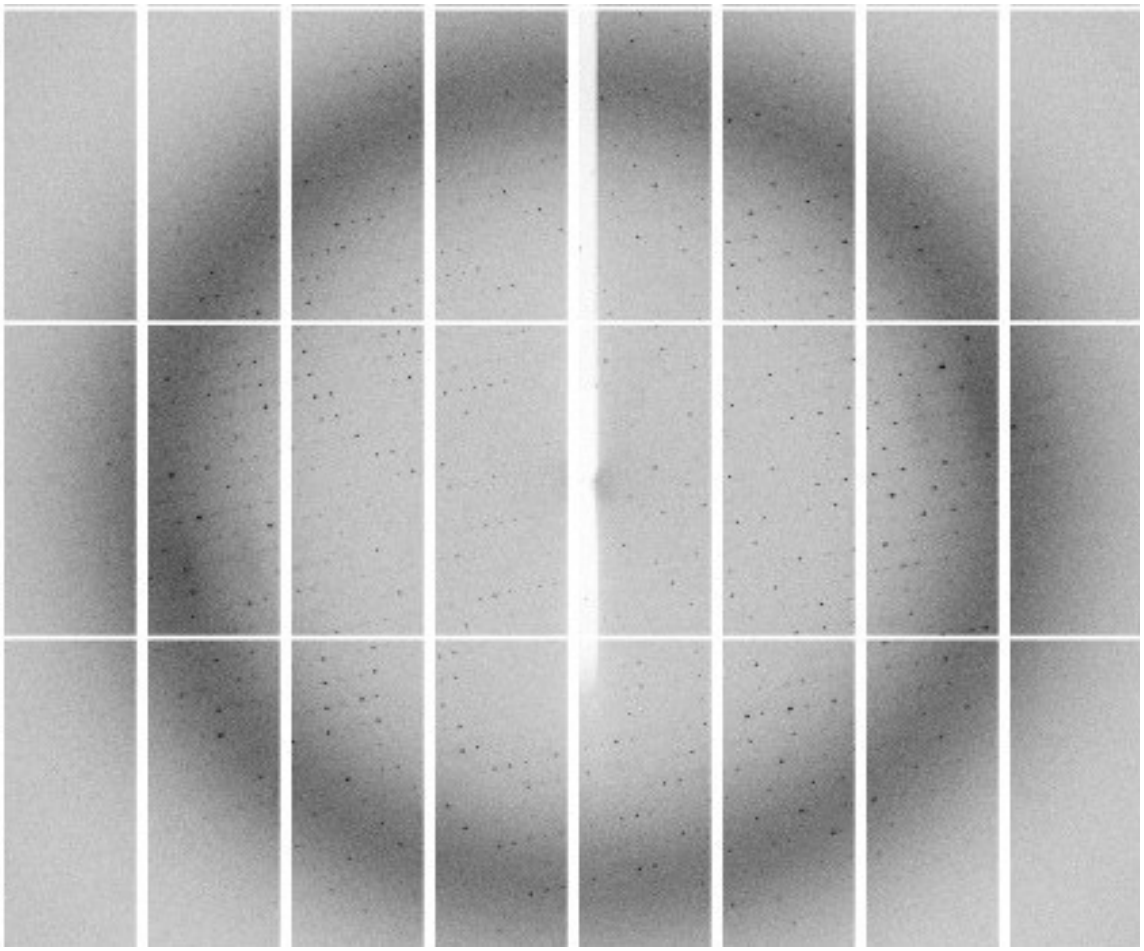


Figure 52: A diffraction pattern of a SAP:Fab crystal at a resolution of 4Å collected at Diamond Light Source Synchrotron, Oxfordshire (more details in Chapter 5). The centre of the film is referred to as the origin and is assigned the coordinates (h,k,l) = (0,0,0). The dark circle seen at 3Å is referred to as the water ring and is caused by the diffraction of ice which is created during the freezing process.

Protein Crystallisation

The first step carried out during x-ray crystallography experiments is crystallisation of the desired protein. Crystallisation is a difficult process to achieve due to the challenges presented by solubilised proteins. Firstly, it is difficult to obtain the high protein concentrations necessary for crystallisation to occur. Secondly, it is common for contaminants to taint samples and thirdly, proteins can be extremely sensitive to temperature, pH and ionic strength. This makes them likely to denature when put through the screening process. Any changes to their environment, even very small ones, could render a protein unsuitable for crystallisation. It is also important to note that every protein is unique in its physiochemical characteristics. This individuality makes optimum crystallisation conditions almost impossible to predict. As a result of this, finding the conditions which make good quality crystals of a particular protein can take months or even years to decipher (Smyth & Martin, 2000) **()**.

In order for crystal growth to occur, non-destructive environmental conditions are required; preferably in the form of physiological conditions. Unfortunately, most proteins are still soluble at physiological conditions. In order to overcome this, proteins call for an increase in solute concentration. This can be achieved by using either salt or precipitants. Each of these is non-disruptive and allow crystal growth. Once formed, the protein crystals are extremely soft, fragile and sensitive to environmental variations. This is due to the fact that they are made of 50-70% solvent and are held together through weak hydrogen bonds, salt bridges and Van der Waals interactions (Smyth & Martin, 2000).

Crystallisation is a form of supersaturation which is often described as a controlled precipitation of protein. If precipitation is not controlled, the protein will leave the solution to form an amorphous precipitate which cannot be used for crystallography. The first stage involved in this controlled precipitation is nucleation. Nucleation involves the formation of well organised aggregates

until a stable nucleus is achieved. This is a rate limiting step and is the cause of the lag phase typically seen in crystal growth curves (sigmoidal). Upon completion of nucleation the growth stage can commence. Here, additional protein molecules are added onto the stable nuclei forming a crystal lattice until limited by external factors such as contaminants, the depletion of protein molecules and the development of growth defects (Ducruix & Giegé, 1992).

It is common to use a phase diagram to provide a visual representation of crystal growth **(Figure 54)**. The solubility curve separates regions which do and do not permit supersaturation (Asherie, 2004). Naturally, in the region which does not allow supersaturation, the protein remains in solution and are therefore useless for x-ray crystallography. The region which does promote supersaturation can allow crystal growth but does not guarantee it. In fact, this supersaturation region can actually be divided into three categories: precipitation, nucleation and metastate. The precipitation region represents conditions in which only amorphous aggregate is formed where as the nucleation region promotes spontaneous nuclei formation and therefore crystal growth. The metastable region, however, represents conditions in which the growth of stable nuclei is promoted but the original formation of those nuclei is inhibited. The metastable region is therefore idyllic for crystal 'seeding' which is a name used to describe when pre-formed nuclei are added to a solution of protein monomers. (McPherson, 1990). Most crystallographers consider it ideal for crystallisation to occur at the borderline between the metastable and nucleation regions. This permits fewer but bigger crystals to form.

In crystallography it is crucial that the protein sample is pure. Contaminants result in non uniform molecules being incorporated into the crystal lattice, causing defects and limiting crystal growth. The pH of the solution is equally important as pH affects the solubility of proteins. Protein solubility is at its lowest just before its pI - the pH at which a protein has no net charge. Also, proteins in low ionic strength solutions exhibit increased interaction between molecules as they attempt to balance their electrostatic charges. Of course, lowering the pH can cause additional problems with protein solubility however

this can be counterbalanced through the addition of salt. The salt acts as a competitor for water molecules and thus lowers protein solubility in a process called “salting out” (Rhodes, 2012).

The addition of precipitants has a similar dehydrating affect to that of salt. Precipitants reduce the dielectric constant of solutions and increase molecular interactions by increasing the strength of electrostatic forces between molecules. Popular precipitants include polyethylene glycol (PEG), ammonium sulfate and 2-Methyl-2,4-pentane-di-ol (MPD) (Ladd & Palmer, 2014).

Proteins with flexible regions in their structure (*e.g.* enzymes) cannot be crystallised in their entirety due to their inability to exhibit the uniform structure necessary for crystal lattices. Fortunately, crystallographers have found two distinct approaches to overcome this. The first of these is fragmentation. This involves the cleavage of the desired protein and individual crystallisation of each fragment. The second approach is called co-crystallisation. Here a super-saturating concentration of substrate, inhibitor or coenzyme specific to the protein is added to the protein solution. The stable complex is separated from the excess components using chromatography methods and then crystallised (Ladd & Palmer, 2014).

Vapour Diffusion

There are various methods in which to form crystals, the most common being vapour diffusion. In vapour diffusion, a droplet containing purified protein, buffer and precipitant is equilibrated with a large reservoir containing similar buffers and precipitants but at higher concentrations. As the two solutions equilibrate the concentration of protein and precipitate in the droplet increase. If the appropriate conditions have been used a crystal will form in the drop. The reason vapour diffusion is one of the more popular choices when attempting crystallisation is because it allows for gentle and gradual changes

in the concentrations. This allows large and well-ordered crystals to form (Chayen, 1998).

There are two ways in which to perform vapour diffusion: the hanging drop method or the sitting drop method. The hanging drop method involves the suspension of the droplet on a siliconised cover slip above a reservoir whereas the sitting drop method involves the placing of the droplet on a pedestal separate from the reservoir (Dessau & Modis, 2011). The hanging drop technique was used predominantly throughout chapter 4 and 5. In this technique, a drop of pure protein is usually mixed with equal volume of crystallisation solution on a siliconised cover slip. This makes the precipitant 50% of the concentration required for crystallisation. The cover slip is placed over a well containing 1ml of the crystallisation solution (drop facing inwards) and sealed with grease so that a closed system is formed. The concentration in the drop is less than that in the well causing vapour diffusion to occur until equilibrium is reached. If the correct conditions have been chosen, supersaturation should occur (Ducruix & Giegé, 1992) **(Figure 56)**.

Robots can be used to set up a large number of crystallisation screens simultaneously. They use the same systems as above but are able to carry out the steps quickly, with a large number of replicates and without human error. Robots are very beneficial as they require only a small amount of protein which allows conservation of protein stocks. Also, the smaller volumes result in quicker crystallisation (Zhu et al., 2014)**.** When making crystals – especially when working with proteins which have never before been crystallised - it is common to use a robot alongside a commercial screen as a preliminary step. In doing so, one is provided with a quick and easy method to determine possible “hits” for crystal growth. Commercial screens use a grid system which varies in reagents, pH, precipitant and salt concentrations. There are multiple manufacturers who produce trays containing 96 pre-made well solutions. Each of the solutions is different allowing a wide range of conditions to be tested in one manoeuvre. Unfortunately, due to the small volumes used, the crystals made from these commercial screens are usually too small for diffraction. Subsequently, the ‘hit’ conditions must be replicated

by a crystallographer on a manual grid. The manual grids can be used to create pH and precipitant gradients around the ‘hit’ as a way to optimise the conditions. It is also common to make changes in protein concentration and temperature in order to achieve optimisation.

Crystal Harvesting

Originally, crystals were harvested by transferring them into a glass capillary tube. A drop of crystallisation buffer was added to keep the crystal hydrated and then the capillary was sealed at both ends, mounted onto a goniometer head and shot with x-rays. This method has both advantages and disadvantages. The advantage of this method is that the crystals are kept in their native well solutions. This prevents them from cracking or drying out. However, the disadvantage is that when shot with x-rays, the crystals which are harvested in this manner are often subject to radiation damage.

Crystallographers have recently discovered the benefit of harvesting and freezing the crystals. The crystals are collected or ‘scooped up’ in loops made from plastic or nylon and frozen. When frozen, crystals experience reduced radiation damage from x-rays due to their increased durability. The free radicals produced during diffraction have a slower diffusion rate and slower mobility in the crystals limiting radiation damage (Elspeth Garman, 2003).

Although there are clear advantages from freezing crystals, initially this proved difficult for crystallographers. They found that the production of ice crystals damaged the lattice and affected results (E. F. Garman & Schneider, 1997). Ice has its own diffraction pattern which can mask the one generated from the protein structure. Nowadays crystals are flash frozen in liquid nitrogen in the presence of an ice preventing cryoprotectant. This prevents the formation of ice whilst still protecting the crystals from radiation damage.

As necessary as cryoprotectants are, they produce their own set of problems. They can cause crystals to crack if an inappropriate one is used or if the

change in solution occurs too quickly. Identifying the correct cryoprotectant requires trial and error which is not ideal as crystals are difficult to obtain. Cryoprotectants are usually a solution of sugar or polyol *e.g.* glycerol and dimethyl sulfoxide (DMSO). They work by forming hydrogen bonds with biological molecules causing the displacement of water molecules. Hydrogen bonds are necessary for proteins to maintain their native state in solution. By displacing the water molecules, the cryoprotectant allows the protein to maintain those hydrogen bonds when frozen without incorporating excess water (Drenth, 2013).

Data Collection / X-ray Exposure

During data collection, the crystal containing loop is mounted onto the goniometer where it is continuously exposed to a stream of nitrogen gas (temp. 100K) to prevent thawing (Drenth, 2013). The goniometer accurately positions the loop within the x-ray beam with the aid of a camera and a computer. This accuracy is essential as both the crystal and the beam are extremely small and are required to be within 25μm of accuracy. There are various types of goniometer with the most common being the kappa goniometer. The kappa goniometer provides 3 angles of rotation (omega, kappa and phi). The omega (ω) angle rotates about an axis perpendicular to the beam. The kappa (κ) angle, which is an axis approximately 50° to the ω axis, allows convenient mounting of the crystal and the phi (φ) angle rotates around the loop/capillary axis. When the κ angle is zero, the ω and φ axes are aligned **(Figure 55)**.

When a crystal is exposed to an x-ray beam it causes the beam to diffract and scatter into a pattern of spots/reflections. The pattern and intensity of each spot is noted by a detector and recorded on a computer. A single picture is not enough to represent the entire Fourier transform therefore the crystal must be rotated step-by-step through 45-180° (depending on the symmetry of the crystal) with images taken throughout in order to collect enough data (Rhodes, 2012).

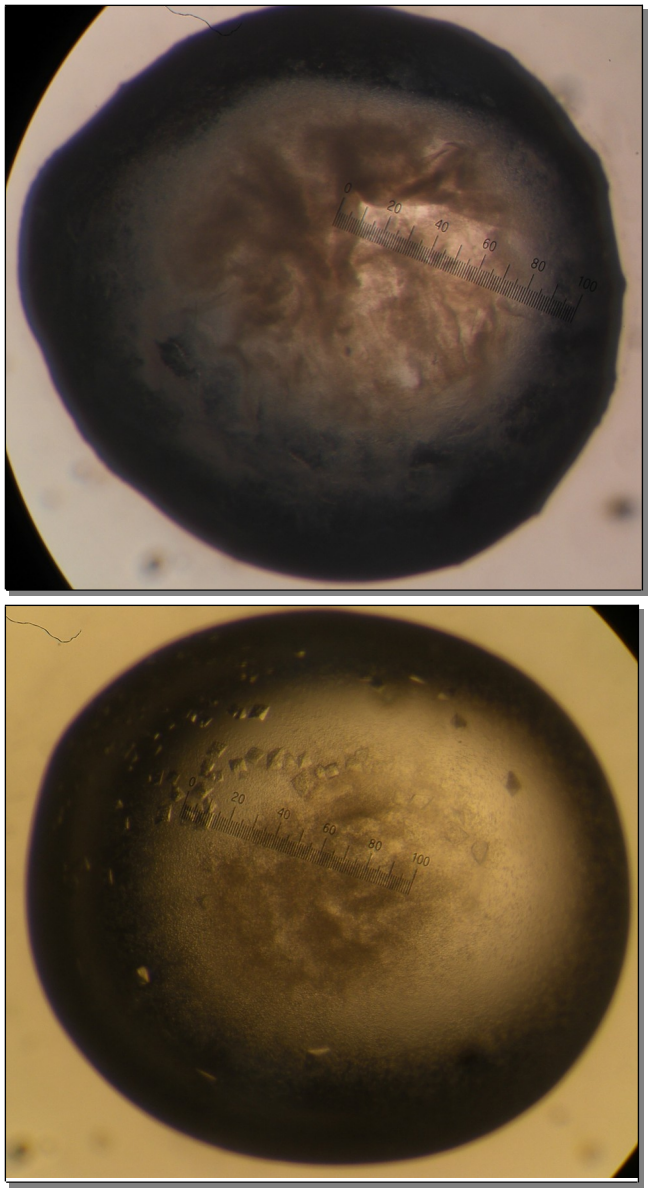


Figure 53: Examples of two SAP crystallisation attempts using the following conditions: 0.1M MgAc, 0.1M MES, 8% PEG10K, pH 5.5. Top – Crystals grown at room temperature resulted in amorphous aggregate. Bottom – Crystals grown at 4°C resulted in multiple square shaped crystals approximately 30µm x 30µm in size.

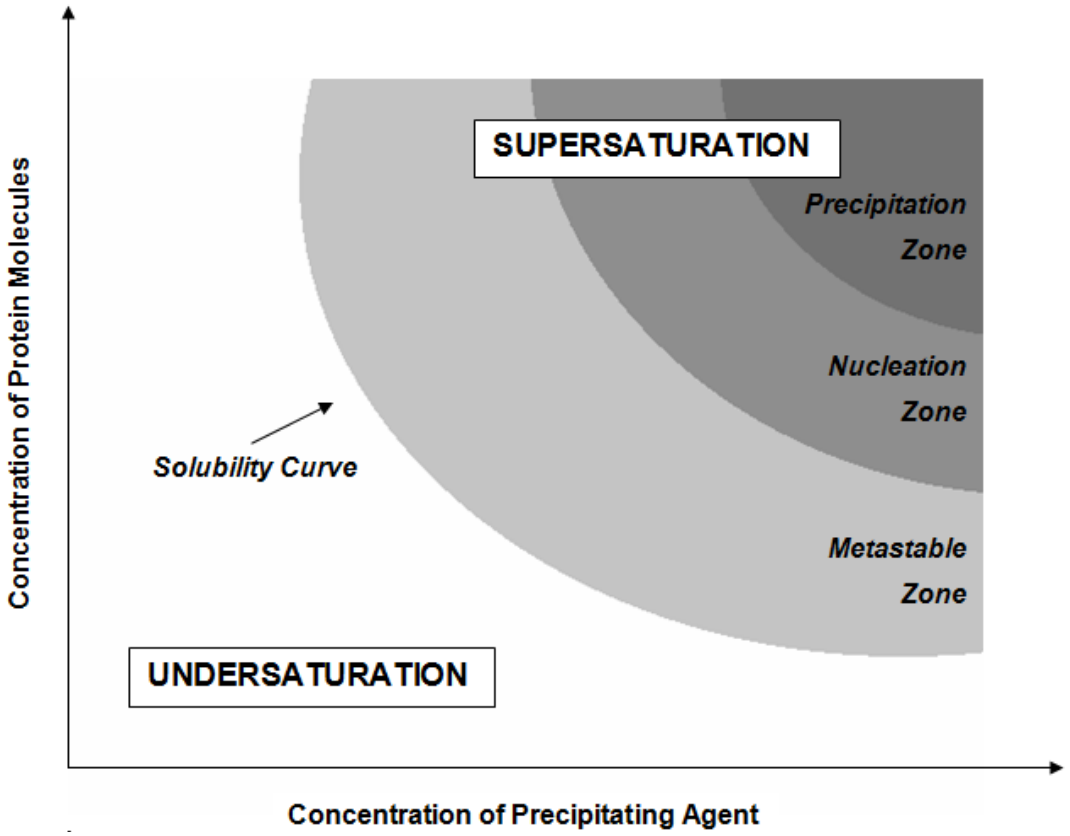


Figure 54: Example of a phase diagram (Ducruix & Giegé, 1992). The solubility curve divides the regions which do/do not allow supersaturation. Of the supersaturation regions, the precipitation zone promotes amorphous aggregate formation. It is the nucleation zone which is responsible for spontaneous nuclei formation and crystal growth whereas the metastable zone promotes growth of stable nuclei but is unsuitable for initial nuclei formation. In crystallisation it is ideal for nucleation to occur between the metastable and nucleation regions as this results in fewer but bigger crystals.

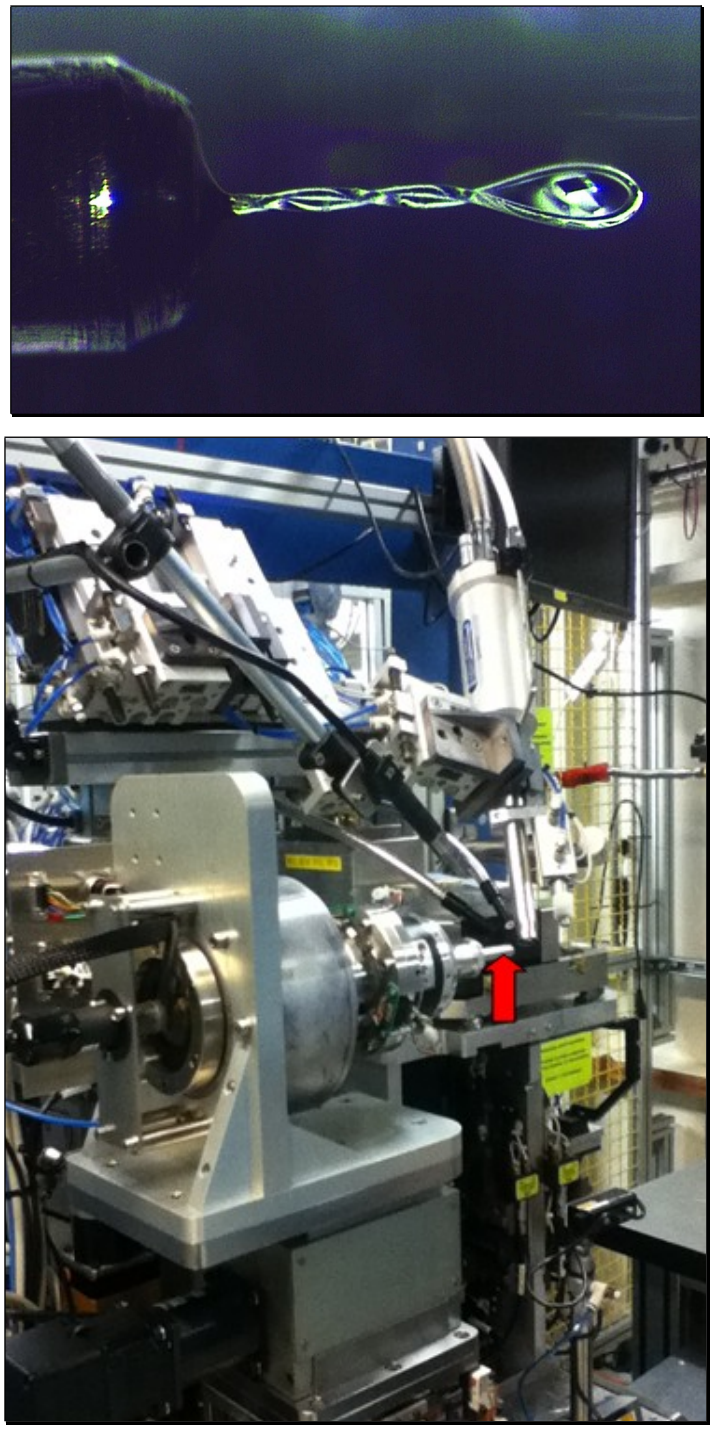


Figure 55: Above – A microscopic view of a crystal suspended in a loop; the crystal is held in place via surface tension of the solution. Below - A photograph of an x-ray machine from Diamond Light Source, Oxfordshire. The crystal containing loops are placed on the goniometer head (red arrow) and subject to a stream of liquid nitrogen to prevent thawing.

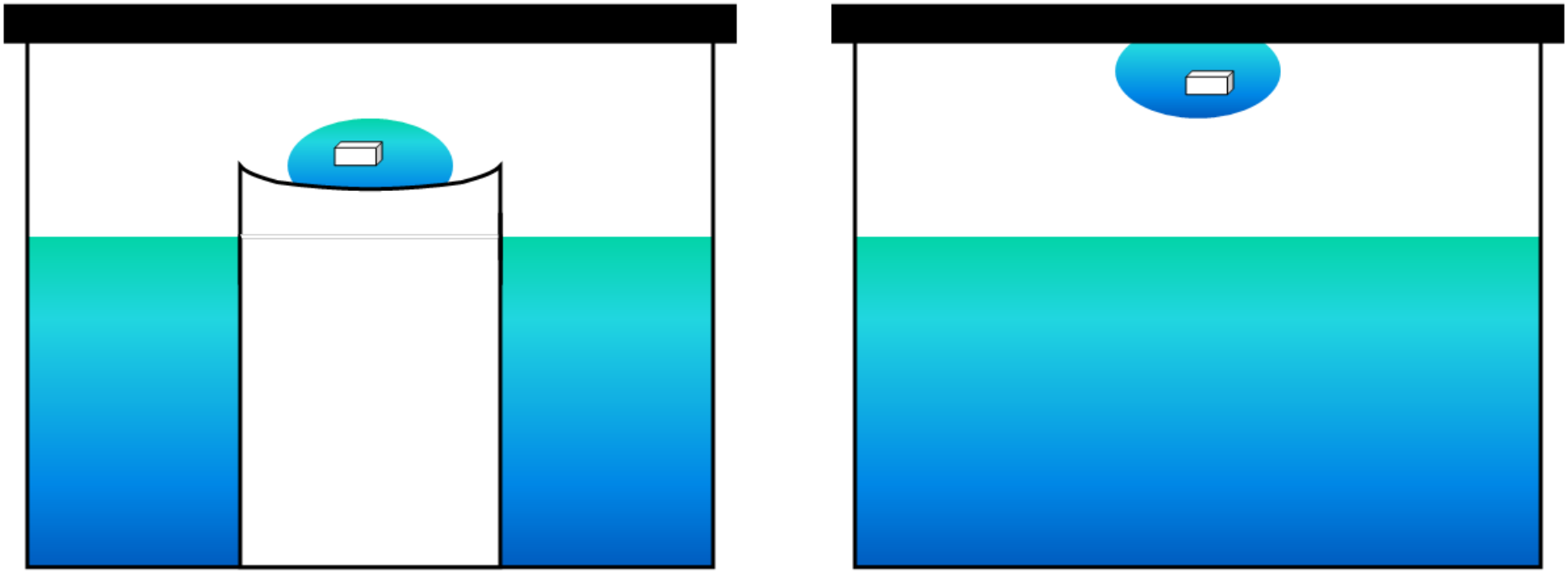


Figure 56: A cartoon diagram to illustrate vapour diffusion techniques. Left – The sitting drop technique. Here the protein containing drop is placed on a pedestal separate from the reservoir well solution. Right – The hanging drop technique. This involves the suspension of the protein droplet on a siliconised cover slip above the reservoir. Both methods are carried out in a closed system to allow vapour diffusion to occur and equilibrium to be reached.

Bragg's Law

In order for an object to diffract a wave, the wave must be no bigger than the object in question. X-rays are a form of electromagnetic radiation with a wavelength of 0.1-100Å. The average bond length is 1.5Å so x-rays in the range of 0.6-3Å are appropriate for atomic structure determination. Light, having a wavelength of 600nm, is unable to interact with the electron cloud surrounding atoms so is useless for this level of structural analysis.

In 1913 William Lawrence Bragg and his father William Henry Bragg discovered that crystalline solids produced surprising patterns of reflected x-rays. They found that crystals produced intense peaks of reflected radiation at certain specific wavelengths and incident angles. They explained this result by modelling the crystal as a set of parallel lines separated by a constant perimeter of *d* as shown in Figure 57. The layers look like rows because the layers are projected onto two dimensions and the viewer's eye-line is parallel to the layers. The variable *d* is the distance between atomic layers in a crystal, and the variable lambda λ is the wavelength of the incident X-ray beam.

The rays of the incident beam are always in phase and parallel up to the point at which the top beam strikes the top layer at atom z. The penetrating X-ray travels down to the internal layer, diffracts, and travels back over the same distance before returning to the surface. The distance travelled depends on the separation of the layers and the angle at which the X-ray entered the material. William Bragg proposed that a diffracted x-ray from the surface of a substance has travelled less distance than the diffracted x-ray from a plane of atoms inside the crystal. If the two beams are to continue travelling adjacent and parallel to one another, this extra distance must be an integral (n) multiple of the wavelength (λ) for the phases of the two beams to be the same. Bragg expressed this in an equation now known as Bragg's Law:

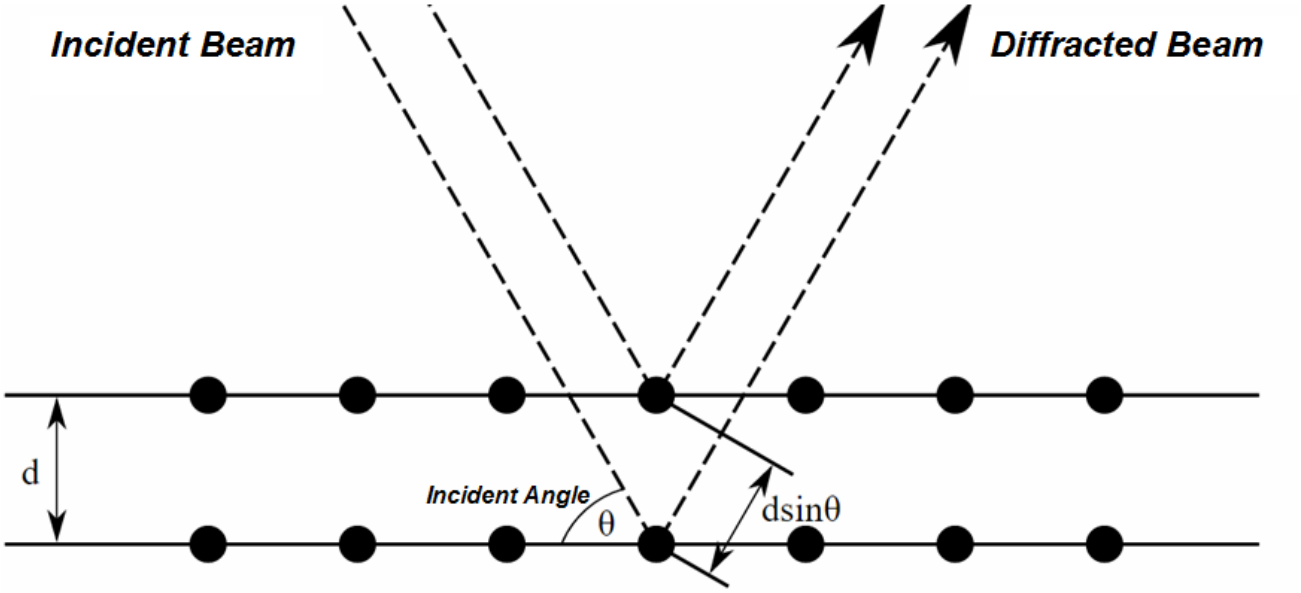
$$n \lambda = 2 d \sin\theta$$

***where* n is an integer and θ is the angle between the incident (or diffracted) ray and the relevant crystal planes**

When **n** is an integer (1, 2, 3 etc.) the reflected waves from different layers are perfectly in phase with each other – *a.k.a*. constructive interference - and produce a dark spot on a diffraction pattern. Otherwise the waves are not in phase – destructive interference - and will either be missing or faint (Bragg & Bragg, 1913).

One of the concluding ideas from Bragg's Law is that diffraction is, in effect, an "arranged event"; three parameters need to be harmonised: the wavelength of the X-rays, λ, the crystal orientation as defined by the angle, θ and the spacing, *d*, of the crystal planes under consideration. For any given wavelength and set of planes, one can conspire to arrange for diffraction to occur by, for example, continuously changing the orientation, i.e. changing theta, until a point arrives when Bragg's Law is satisfied. This is precisely when diffraction occurs.

Bragg's law confirmed the existence of real particles at the atomic scale, as well as providing a powerful new tool for studying crystals in the form of X-ray and neutron diffraction. In fact, for this William Lawrence Bragg and his father, Sir William Henry Bragg, were awarded the Nobel Prize in physics in 1915 for their work in determining crystal structures.



A diagram to illustrate Braggs Law:
d = distance between atomic planes
θ = angle of incidence/diffraction.

Figure 57: Braggs Law states that the x-ray wave which is diffracted from the surface of a substance travels less distance than the x-ray which is diffracted from a plane of atoms inside the crystal. The extra distance travelled by the penetrating wave depends on the separation of the layers and the angle at which the X-ray entered the material. When the reflected waves from different layers are perfectly in phase with each a dark spot is observed on the diffraction pattern. When the waves are not in phase, spots will either be missing or faint.

X-Ray Sources and Detectors

There are three forms of x-ray source: x-ray tubes, rotating anodes and particle storage rings. The tubes and anodes generate x-rays by bombarding a metallic target – normally copper – with a beam of electrons. Following collision, an electron from a low lying orbital in the target metal atom is displaced creating a vacancy in its orbital. This is then filled with an electron from a higher orbital. As the electron drops, excess energy is released in the form of an x-ray photon (Drenth, 2013) **(Figure 58)**.

In the case of particle storage rings – *a.k.a.* synchrotrons - electrons are sped up to 99% of the speed of light. They are contained in vacuum storage rings, sustained via magnets and driven in a circular motion by energy from radio frequency transmitters. The sustaining of the electrons in this circular motion produces x-ray beams which are deflected from side to side using magnets. This creates a fan of x-rays which can be used in several workstations located at regular intervals around the ring. The x-ray beam is optimised prior to diffraction via monochromators and mirrors. The mirrors focus the beam onto the crystal whereas the monochromator selects the required wavelength from the spectrum and remove unwanted radiation (Whittaker, 1982). Monochromators are made from crystals of silicon, germanium or carbon each of which will select a wavelength upon diffraction (Drenth, 2013).

Partical storage rings (synchrotrons) are the preferred method for creating x-rays as the resulting beam is more powerful and intense compared to the alternative methods thus produces higher quality data at a higher resolution (Lindley, 1999). They have reduced exposure times from hours to minutes causing a reduction in crystal damage, and they also collect an increased amount of data. One further advantage of using storage rings is the fact that the x-rays produced are tuneable. This quality proves useful in Multi and Single wavelength anomalous diffraction (MAD/SAD) where wavelength specificity is used to solve the ‘Phase Problem’ – a term used to describe the loss of information concerning the phase of the x-ray wave during

measurement (more details to follow). By collecting data at several wavelengths one can determine phase information by measuring the difference in anomalous scattering (Drenth, 2013).

Although synchrotrons have proved to be the superior method in diffraction, ‘in house’ x-ray sources are still in use today. ‘In house’ sources are much cheaper and easier to maintain due to the x-ray source being produced by a rotating copper anode rather than vacuumed electrons. Unfortunately the repetition rate of the ‘in house’ data collection laser is limited to a few hertz rather than the megahertz repetition rates naturally arising in storage ring emission. This means crystals have to undergo radiation for much longer in order to produce the same amount of images risking damage to the crystal. However, ‘in house’ does has the advantage of convenience which is why it is still used as a screening method today. Diamond Light Source Synchrotron is one of two synchrotrons in Europe and thus requires travel and shift work in order to use it (Skarzynski, 2013) **(Figure 59)**.

Various detectors can be used during x-ray crystallography however the most common is the charged coupled devices (CCDs). CCDs are essentially photon counters. They are covered with phosphorus and emit visible light when exposed to diffracted x-rays. When an entire diffraction pattern is collected the accumulated pixels are converted into a digital signal for subsequent processing by a computer (Drenth, 2013).

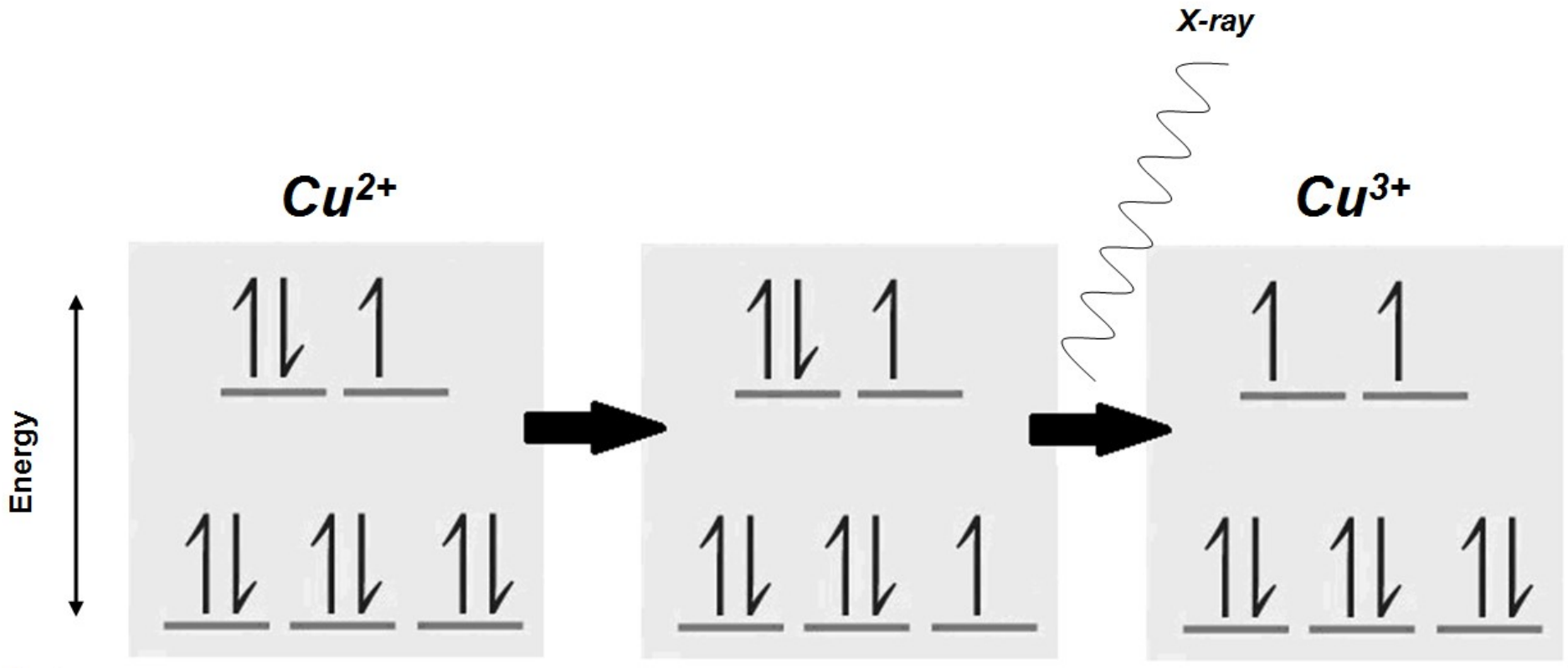


Figure 58: A diagram to demonstrate x-ray generation in tubes and anodes. Initially there is a collision of copper (Cu^{2+}) with a beam of electrons. This causes the displacement of an electron from a low lying copper orbital. This is subsequently replaced with an electron from a higher orbital causing excess energy to be released in the form of an x-ray photon.

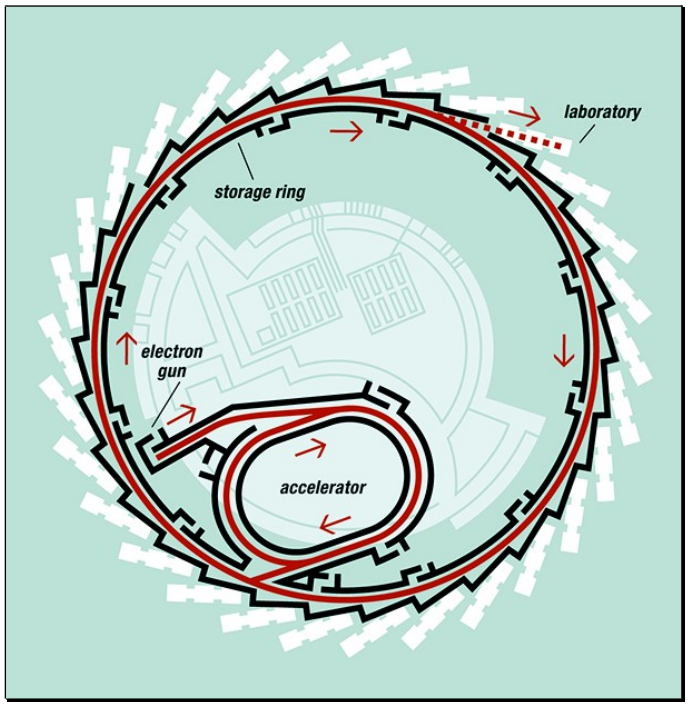


Figure 59: Photographs of Diamond Light Source, Oxfordshire. Above – A cartoon of the interior structure. The accelerator is located in the centre of the ring and is responsible for accelerating electrons to 99% the speed of light. Once accelerated, they are contained in the vacuum storage ring and sustained via magnets in a circular motion. The sustaining of the electrons in this circular motion produces x-ray beams which are used in several workstations located at regular intervals around the ring. Below – A digital photograph of the exterior structure.

Data Processing

When the crystal is hit by x-rays, the beams diffract at unique angles depending on which atom they make contact with. This generates a spot on the diffraction pattern with particular co-ordinates. The coordinates of each spot is measured in relation to the centre so to give it a value; the centre of the diffraction pattern has the coordinates (h,k,l*)* = (0,0,0). Data processing is required in order to turn the raw experimental data into a manageable data set in the form of an .mtz file. The .mtz file contains the Miller indices (hkl), intensities (I_{hkl}) and standard uncertainties (σI_{hkl}) of all the reflections in the diffraction experiment that are required for the Fourier transform formula below. Through this, the electron density surrounding an average molecule in the unit cell ($\rho_{x,y,z}$) can be described.

$$\rho(x,y,z) = \frac{1}{V}\sum_{h}\sum_{k}\sum_{l} F_{hkl}\, e^{-2\pi i(hx+ky+lz)}$$

There are three stages in data processing: autoindexing, post-refinement and integration. During the first stage – autoindexing – the unit cell parameters are determined alongside the point group and orientation of the crystal in relation to the x-ray beam. The unit cell is the most basic repeating structure of a lattice. It has three lengths (a,b, and c), three angles (α β and ɣ) and can be divided into asymmetric units. The values of these dimensions describe the morphology of the unit cell **(Figure 61)**.

The symmetry of a unit cell is described by its space group. The beginning letter of a space group represents the lattice type and the numbers represent the operations defining the symmetry of the cell. (*e.g.* $P2_12_12_1$) There are 5 lattice types: P, C, F, I, and R. P (primitive) is the smallest lattice type and the most desired as it contains a single lattice point at each corner around which the asymmetric units are arranged. The other lattice types are said to be face-

centered or body-centered as they contain (in addition to the lattice points at the corner of the cell) extra points in the centre of one face (C), all faces (F) or the centre of the cell (I). The unit cell can adopt one of seven different crystal systems: triclinic, monoclinic, orthorhombic, trigonal, tetragonal, hexagonal or cubic. A combination of these seven crystal systems plus the 5 types of centering allows the possibility of 14 Bravias lattices (Ladd & Palmer, 2014) **(Figure 62)**.

The post refinement stage follows autoindexing. This process refines the initial estimates of the cell parameters by combining one or two segments of data, each containing two or three different images. Following this, the final integration stage merges all of the collected images and refines the crystal and detector parameters to generate a full list of hkl and I_{hkl}.

Upon further scrutiny one will notice that the actual diffractors of x-ray beams are the clouds of electrons around the atom, not the atom itself. – *i.e.* diffraction patterns represent electron distribution/density surrounding a molecule. It is the crystallographers' job to rearrange the atoms so that they fit within this electron density map.

The Phase Problem

The Fourier transform above is used to convert diffraction data into the electron density of the unit cell. Most of the parameters needed to construct this equation are available from the diffraction data. A diffraction wave is made up of 3 factors; frequency, phase and amplitude **(Figure 60)**. The frequency is that of the x-ray source and the amplitude is equal to the square root of the intensity measured for the reflection hkl $(I_{hkl})^{1/2}$. The phase of the wave, however, is unknown. This loss of phase information is known as the 'Phase Problem'.

There are three methods used to overcome the 'phase problem'

- Isomorphous replacement

- Single/multi-wavelength anomalous dispersion (SAD/MAD)
- Molecular replacement

Isomorphous replacement involves the insertion of a heavy atom into the structure either by crystal soaking or through protein synthesis with analogues. These heavy atoms are used to help orient and position the reflections. The advantage of using this method is that the addition of the heavy atom does not alter the crystal formation or the unit cell in any way; it is isomorphic to the native crystal. The disadvantage is that at least two isomorphous derivatives must be evaluated - the native and heavy - since using only one will give two possible phases. This is time consuming and can use large quantities of protein sample. However, once both data sets are collected the Patterson difference map can then identify the location of the heavy atom in the map determining both amplitude and phase.

MAD is when diffraction is recorded at multiple wavelengths. The atoms inner electrons absorb x-rays and reemit them after a delay causing a phase shift in all of the reflections. This is known as the anomalous dispersion effect. By analyzing the phase shift rather than just the individual reflections, a solution for the phases can be provided. Here, the advantage is that the method does not require multiple crystal structures with heavy and light atoms. The difficulty with this method is that it requires very specific wavelengths which can only be obtained using synchrotron radiation.

In the case of SAD, a single dataset is collected at a single wavelength. The advantage of this method is the reduction of beam exposure to the crystal, minimising radiation damage during data collection. The disadvantage lies in the quality of the data produced. The electron density map of the resulting data must be adjusted at the interstitial regions between protein molecules occupied by the solvent in a process called ‘solvent flattening’. Only then can the results be comparable with those derived from MAD experiments. This is a time consuming process.

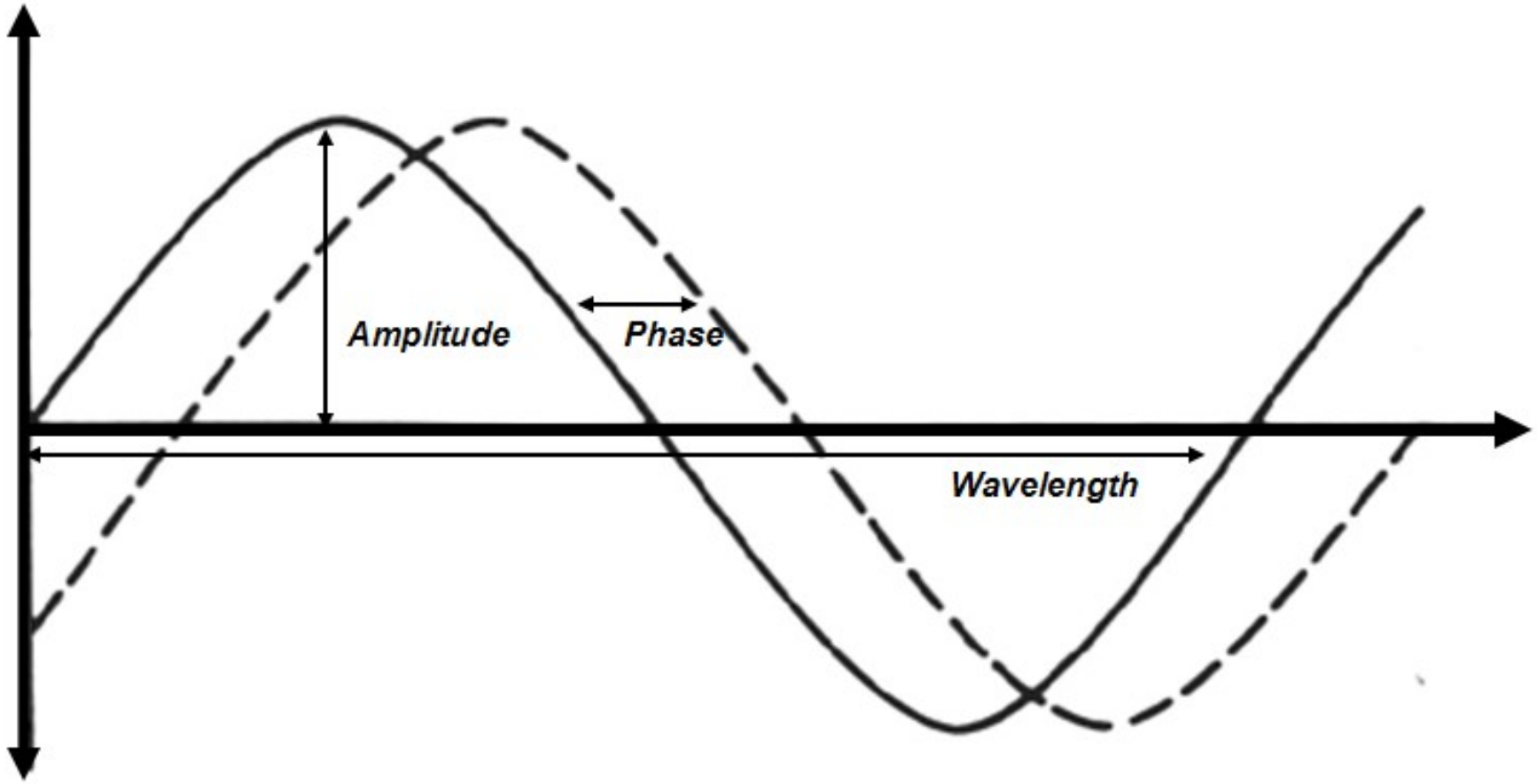


Figure 60: A graph to illustrate the features of an x-ray wave. Amplitude represents the maximum displacement of a point in a vibrating wave whereas wavelength is the distance between successive crests in a wave. To explain phase, one must remember that waves are composed of peaks and valleys with a zero-crossing between them. The phase of a wave is described as the distance between the first zero-crossing and the point of origin.

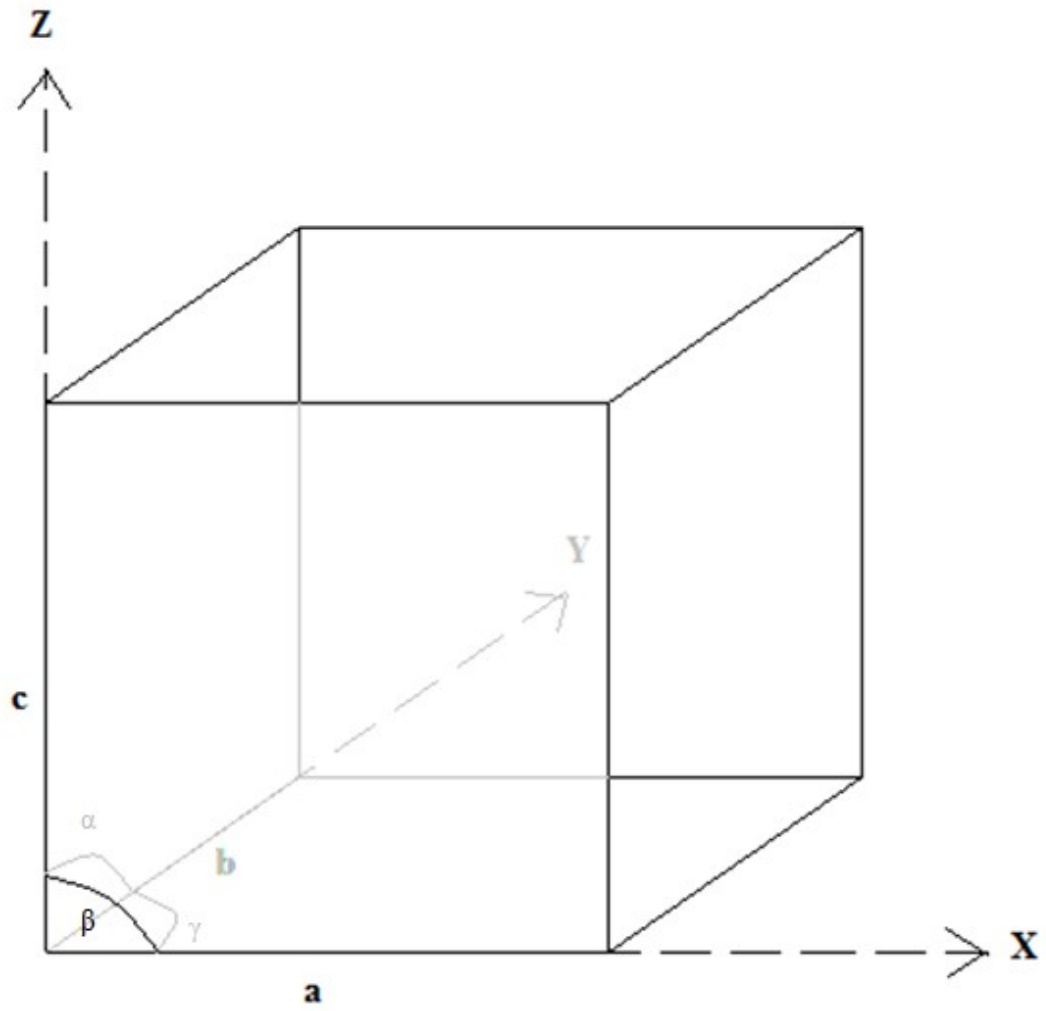


Figure 61: The dimensions of a unit cell.

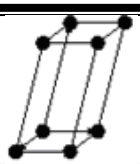
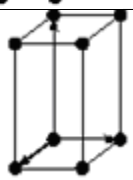
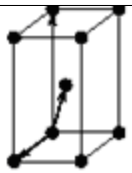
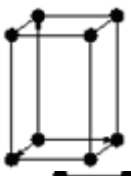
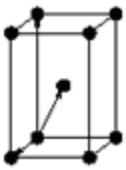
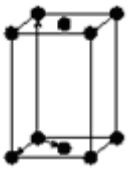
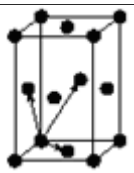
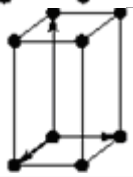
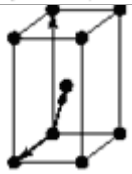
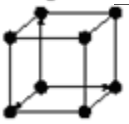
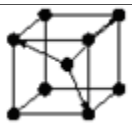
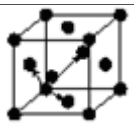
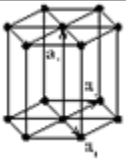
Bravias Lattice	Parameters	Simple (P)	Volume Centered (I)	Base Centered (C)	Face Centered (F)
Triclinic	$a_1 \neq a_2 \neq a_3$ $a_{12} \neq a_{23} \neq a_{31}$				
Monoclinic	$a_1 \neq a_2 \neq a_3$ $a_{23} = a_{31} = 90°$ $a_{12} \neq 90°$				
Orthorhombic	$a_1 \neq a_2 \neq a_3$ $a_{12} = a_{23} = a_{31} = 90°$				
Tetragonal	$a_1 = a_2 \neq a_3$ $a_{12} = a_{23} = a_{31} = 90°$				
Trigonal	$a_1 = a_2 = a_3$ $a_{12} = a_{23} = a_{31} < 120°$				
Cubic	$a_1 = a_2 = a_3$ $a_{12} = a_{23} = a_{31} = 90°$				
Hexagonal	$a_1 = a_2 \neq a_3$ $a_{12} = 120°$ $a_{23} = a_{31} = 90°$				

Figure 62: A table of the 14 Bravias lattices.

Molecular Replacement

Molecular replacement was the chosen method for this project. This is where the phases of a similar, previously solved structure are used to orient the experimental molecule in the unit cell. The advantage of this is that the calculated phases can be applied to the experimental molecule in order to solve the structure. The disadvantage of this method is that it can severely bias the solution of the experimental structure. In order to use the molecular replacement method there needs to be homology of at least 25% between the model and experimental structures. To prevent model bias from occurring the structures are put through multiple cycles of refinement, map calculation and model building.

Molecular replacement requires a six dimensional search which includes three translations and three rotations. By splitting this search into two different stages efficiency is maintained and computing time is minimized. The model is first rotated whilst ignoring its symmetry until the best agreement with the observed data is observed. The correctly orientated structure is then translated within its parameters until best agreement with the diffraction data is obtained.

The 'R factor' is used to evaluate the translation function. It monitors the relationship between the model structure amplitudes (calculated) and the experimental ones (observed).

$$R = \frac{\Sigma\left[\left|F_{obs}\right|\right] - \Sigma\left[\left|F_{calc}\right|\right]}{\Sigma\left[\left|F_{obs}\right|\right]}$$

With each reflection the difference between the observed and actual structure factors are calculated and the magnitudes of the difference for each reflection are added together and divided by the sum of $F_{obs.}$ If the agreement is close then the sum of the differences will be smaller than the intensities themselves

and R is therefore small. A perfect agreement will have an R factor of 0 but crystallographers consider values between 0.3 and 0.4 to be respectable. The final value of the R factor never actually becomes zero as its value depends on the degree of order held within the crystal and the quality of the diffraction data; neither of which will be perfect.

Although an extremely useful technique, molecular replacement is only a preliminary method of structure solving. In order to refine the model, rigid body refinement must be performed.

Structure Refinement

Upon solving the 'Phase Problem' using one of the three techniques mentioned above, the data must continue to be processed with structural refinement and modeling techniques. The purpose of refinement is to achieve the best fit between the pre-solved model (calculated data) and the experimental data (observed data). This is done by altering the atomic positions x,y and z until an improved fit is obtained. Again, the R factor is used to measure the discrepancies between the calculated and observed data. As agreement between the two improves, the R factor falls down towards zero.

The refinement of proteins is much more difficult than the refinement of their inorganic counterparts. This is due to the large quantities of atoms present in a protein structure. Each atom has three parameters which need to be included in the refinement process: occupancy, temperature factor (B-factor) and atomic position. The more atoms there are, the weaker the ratio between observations and atomic parameters. To resolve this ratio problem, a number of constraints and restraints are used in the refinement process.

A constraint is a fixed value for a certain parameter. An example of this is when the occupancy of a particular atom is set to 0.1 during refinement. A

restraint allows slightly more flexibility in value but in this case bond lengths and angles must be within strict range of values. The inclusion of constraints and restraints in refinement acts as an equivalent to increasing the number of observations. It also helps to maintain the correct stereochemistry of the model.

There is a limit to the amount of constraints and restraints that can be used in refinement. This is dependent on data quality and resolution. The lower the resolution – *i.e.* the lower the amount of reflections – the fewer observations there are ergo limiting the ability to refine all the atomic parameters individually. The observations available at relatively low resolutions would only permit the refinement of the main chain conformational angles (torsion angles). This generates two variables per amino acid: φ and ψ. In the case of higher resolution data, side chain conformation angles (χ) can also be refined but only with the bond angles, bond lengths and planarities tightly restrained with a single B factor for the whole structure. This circumstance provides four variables per amino acid. With further improvement of the data two separate B factors can be calculated for each amino acid (one for the main chains and another for the side chains) generating three more variables per amino acid. Resolutions of 2.5 Å or better allow separate B factors to be calculated for each atom. At 2.0 Å, cartesian angles can be refined for each atom. Upon reaching 1.5 Å, anisotropic B factors showing density contribution of individual atoms and the particular direction in which they are free to move, are added as parameters.

Model Building

As the name suggests, model building is used to rebuild the structure model into the electron density map wherever necessary. This is done manually by the crystallographer and can be considered as the final “polishing” step in structure solving.

This method requires the electron density to be calculated, $\rho_{(x,y,z)}$. Normally two types of electron density map are calculated: a $2F_o$-F_c map and a F_o-F_c map. The F_o-F_c map provides information on the new model. The resulting Fourier term can be either positive or negative depending on whether F_o or F_c is larger. A positive Fourier term indicates that the observed intensities (F_o) to ρ are greater than that of the calculated structure factors (F_c). This in turn suggests that the unit cell contains more electron density in this region than implied by the model (F_c) and requires the movement of atoms towards this area in order to increase the electron density. A negative Fourier term indicates that the model implies more electron density in the region than the unit cell actually contains. Atoms must therefore move away from this area **(Figure 63)**.

The F_o-F_c map can be difficult to interpret on its own so the more interpretable $2F_o$-F_c map is also calculated. By doing so, the influence of F_c is still reduced but to a lesser extent than the F_o-F_c map. If errors in the model are minimal, this map is completely positive and depicts a molecular surface covering all the atoms of the protein.

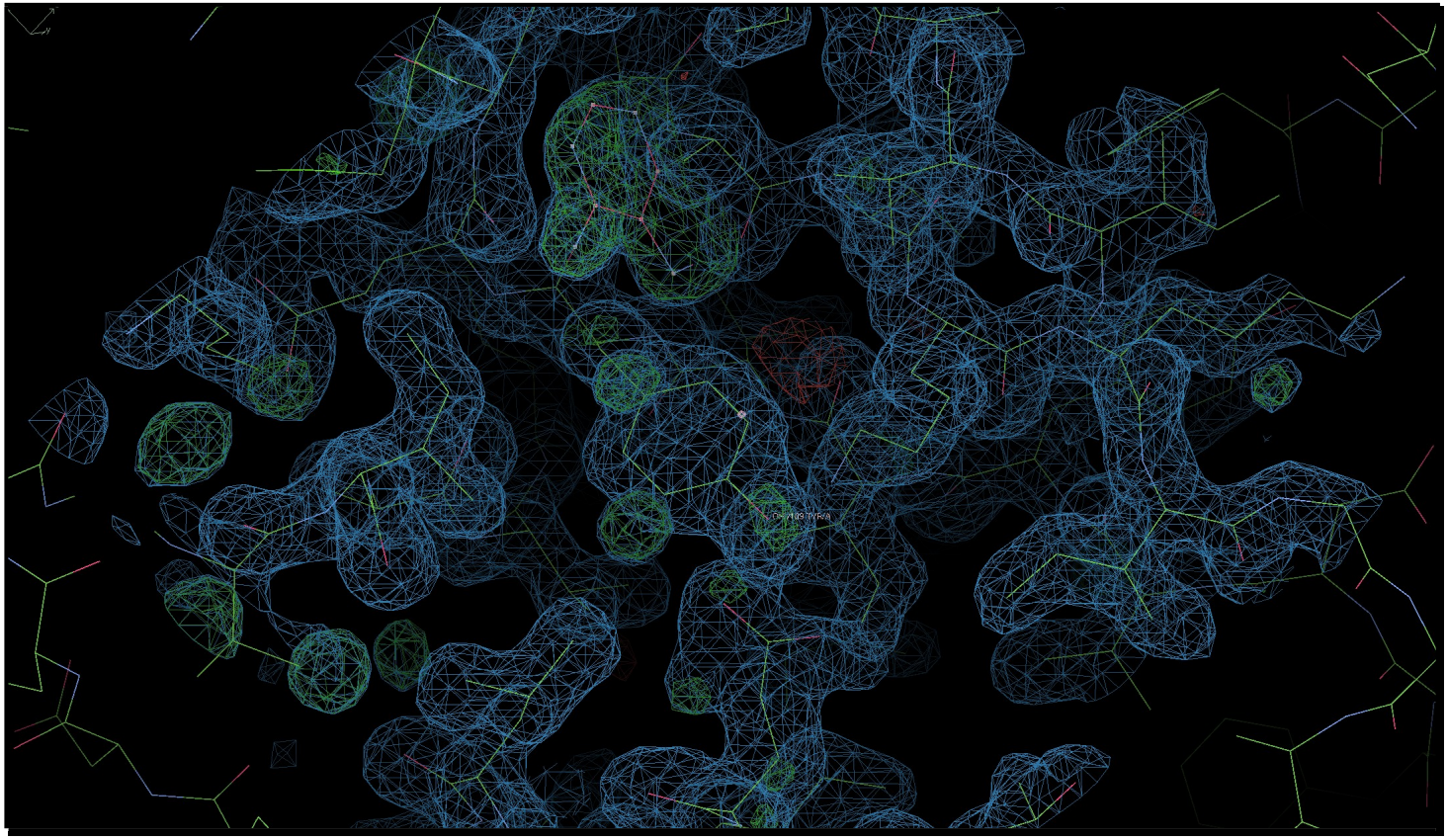


Figure 63: An example of a model constructed using an electron density map. The protein chain is represented in yellow sticks and the blue netting represents the clouds of electrons surrounding the atoms, not the atoms themselves. The atoms are manipulated manually so to fit them into the density map in a process called 'modeling'.

Chapter 5

Amyloid Treatment and Structural Studies

Antibodies against SAP

Chapter Aims

This chapter aims to look at the relationship between SAP and its specific antibody through use of x-ray crystallographic techniques. The reason for this refers back to O'Nuallain and Wetzel's work on SAP knockout mice in 2002. They made the following conclusion:

'SAP has an anti-opsonin role' - *i.e.* the immune system is normally able to recognise amyloid but SAP provides a protective coat around the fibrils. This in turn prevents phagocytic cell recognition and thus prevents an immune response in the presence of amyloid deposits.

Chapter 4 briefly mentioned GlaxoSmithKline's recent developments on antibodies specific to SAP. These anti-SAP antibodies have the ability to detach SAP from amyloid deposits and thus leave the deposits exposed to the host's immune system. The exact mechanism behind this is not fully understood however clinical trials have shown that following the removal of deposit bound SAP from the body, deposits are dissociated and fibre mass is reduced.

Unfortunately, affinity towards the anti-SAP antibody is not specific to deposit bound SAP alone. Serum SAP also exhibits optimal binding to the antibody. Due to this, an anti-SAP antibody dosage often exhibits complete saturation before it has even been able to reach its deposit bound target. In an attempt to resolve this problem, the drug R-1-[6-[R-2-carboxy-pyrrolidin-1-yl]-6-oxo-hexanoyl] pyrrolidine-2-carboxylic acid (CPHPC) has very recently been used in conjunction with the anti-SAP antibody in clinical trials.

CPHPC acts as a competitive inhibitor of SAP. CPHPC is a small palindromic molecule which has the ability to remove 99% of SAP from circulation within a

single week of administration. X-ray crystallography has been used to better understand the mechanism in which it does this. The resulting structure shows that five CPHPC molecules dimerise two SAP molecules, B-face-to-B-face, via cross-linking of lysine 143 in neighboring SAP monomers **(Figure 64)**. When CPHPC is administered in patients, the resulting decameric complex is immediately removed from the system via the kidneys and liver and as of yet, no side effects have been noted. Once treatment has ceased, serum levels are restored to their original concentrations after four weeks (Pepys et al. 2002). It has also been noted that CPHPC has the ability to detach and remove *deposit* bound SAP from the system in a fashion similar to the antibodies. However, this is done inefficiently which explains why it has not been used as a lone treatment process until recently.

In order to efficiently dissociate amyloid deposits it is necessary to use both anti-SAP antibodies and CPHPC in conjunction with one another. Treatment begins when CPHPC is introduced to the system. The drug readily clears the circulatory SAP from the system leaving the deposit bound SAP. Following this, the anti-SAP antibodies are administered. These antibodies bind to the remaining SAP - those which re bound to the amyloid deposits - and cause them to dissociate from the fibres thus leaving them exposed. This stimulates an immunological cascade. Signals created by the antibodies activate a compliment dependent, macrophage-derived, giant cell attack on the newly exposed fibres and the deposits are depleted (O'Nuallain and Wetzel 2002).

The use of this combined treatment against amyloid originally proved successful in 2010 when Bodin *et al* attempted this on mice carrying visceral amyloid deposits (Bodin et al. 2010). They found that upon addition of the antibody, amyloid deposit mass was rapidly reduced with no clinical or biochemical adverse effects. Currently, anti-SAP antibodies are being used in conjunction with CPHPC as a clinical treatment for human patients suffering from amyloidosis at the Royal free Hospital, London. The latest results show that the treatment appears to be safe, reduces amyloid in almost all subjects and can even improve organ function. The depletion of amyloid is not equal in all patients but the clearance of amyloid is much faster than that with other

treatments. The most recent trial was conducted on a total of 16 patients ranging between 18 to 70 years of age. Upon administration of CPHPC, circulating SAP was depleted to less than 0.5 mg per litre in patients with a small or moderate amyloid load and to less than 2 mg per litre in patients with a large amyloid load. Following treatment with anti-SAP antibodies, only transient low-grade infusional side effects were observed. Three patients experienced brief episodes of headache and another a brief episode of nausea (Richards et al. 2015).

Although these results appear promising, more work needs to be carried out in order to guarantee safety and effectiveness of the treatment. Thus far these trials have shown a decrease in amyloid but they have yet to be successful in complete amyloid eradication. It is necessary to remove the entire deposit otherwise symptoms may return and surgery may still be required at a later date. By understanding the location and mechanism behind SAP:antibody binding it may become apparent how to optimise the treatment process.

To this date, extensive structural analysis has not been carried out on the SAP:antibody complex. It is not known how or where they bind. This limits the understanding scientists have on the interactions between the two molecules and how detachment from *in vivo* fibres actually occurs. In order to add some clarity towards this, the crystal structure of SAP bound to its specific antibody has to be made and subject to x-ray diffraction analysis. Attempts to carry out such a protocol have never been done before, however, it should be possible if the following steps are carried out:

- Digest the SAP specific antibody so to remove its flexible hinge region and thus make it suitable for crystallisation.
- Combine SAP with an excess of the digested fragments in order to create a fully saturated SAP:Fab complex.
- Decipher the optimal conditions required for crystallisation of the complex.

- Submit the complex crystal to x-ray analysis in order to obtain a diffraction data set.
- Solve the diffraction data set using the CCP4i programme suite so to decipher the interaction between the SAP and Fab fragments.

By solving the structure of this complex, the mechanism behind SAP:antibody binding could be distinguished. This could lead to later developments regarding this treatment thus allowing for higher success rates and lower fatality numbers.

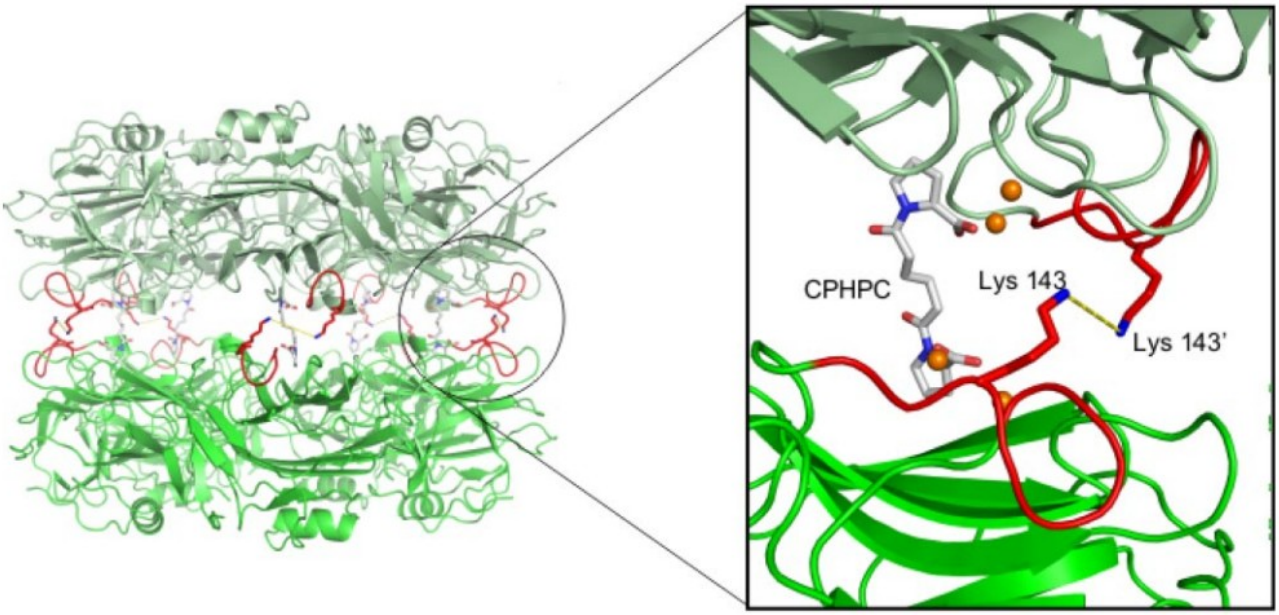


Figure 64: Images depicting the interaction of SAP with CPHPC. Above – A stick diagram showing the point of cross-linking in the SAP sequence. Below – A structure of the complex in solution. Two SAP pentamers (green) are cross-linked, B-face to B-face, by five molecules of CPHPC via lysine143. The residues involved in binding (138 – 144) are shown in red and calcium ions are represented in orange. (PDB code: 4AVT) (Kolstoe et al. 2009)

An Introduction to Antibodies

The body is continuously exposed to a variety of foreign materials (antigens). These antigens must each be eliminated quickly and efficiently to prevent any harm to the body. It would not be practical for nature to devise a different molecular solution for each antigen as the genome is not large enough to contain that amount of information. Instead, nature has created a single way in which to resolve the antigen problem; the antibody.

An antibody has two functions in the immune system. The first is to recognise and bind antigens whereas the second is to trigger the elimination of that material via effector molecule binding. The first of the two functions requires a great deal of diversity as the antibody has to be able to recognise and bind a divert number of molecular structures. The second function, however, requires a certain level of consistency in order to trigger the appropriate immune response upon binding. The conflicting requirements are met by the antibody structure (Delves et al. 2006).

The antibody is 'Y' shaped and consists of three units. Two of these units are identical and are involved in antigen binding. These are called the fragment antigen binding units (Fab). The Fab units contain protein sequence segments which vary greatly between antibodies and provide binding specificity to a large number of antigens. The remaining fragment crystalline unit (Fc) has little variation in its primary protein sequence and is responsible for triggering the immune system through effector molecule binding (Janeway et al. 2001).

There is an extended sequence segment which connects the Fab and Fc regions referred to as the hinge region. The hinge region adds flexibility to the Y structure allowing the antibody to bind to molecular structures of different spacing on an antigen cell surface **(Figure 65)**. Although this flexibility is useful for molecular function, it is an undesirable quality in terms of x-ray crystallography. Fortunately the hinge region is extremely susceptible to

cleavage by proteases which enables non flexible (*i.e.* uniform) structures to be produced (Delves et al. 2006).

There are five antibody subclasses found in humans named Immunoglobulin γ, α, μ, δ and ε (IgG, IgA, IgM, IgD and IgE). Each is responsible for binding a different effector molecule and thus triggers a different pathogenic system upon antigen binding. All five antibody subclasses are composed of a four-chain structure which consists of two identical 'heavy' and two identical 'light' polypeptide chains. The heavy chains span the Fab and Fc units whereas the light chains are associated purely with the Fab unit. The five subclasses are mostly identical however they each have variations in their Fc region which is what determines their unique trigger system. The structural variety of each subclass also determines the polymerisation state of the monomer. For example, IgM predominantly exists as a pentameric form where as the other subclasses are monomeric. The IgA subclass does have the ability to form a dimer but only exhibits this when excreted in seromucous secretions (Delves et al. 2006) (Schroeder Jr. and Cavacini 2010) **(Table 7)**.

Immunoglobulin G (IgG) Structure and function

IgG is the most abundant immunoglobulin form. It is synthesised and secreted by plasma B cells and exists up to 75% in serum (Hadley 2007). The heavy chains in IgG are approximately 50kDa in length while the light chains are half the size at 25kDa each. This totals to 150kDa per antibody molecule. The heavy chains are linked together via two inter-peptide disulphide bonds and each has a light chain attached via another inter-peptide disulphide bond (Schroeder Jr. and Cavacini 2010).

There are 4 IgG subclasses in humans (IgG1,2,3 and 4) named in order of abundance with 1 being the most abundant **(Figure 66)**. They have a sequence homology of 95% with differences lying in their hinge region. Of the four, IgG1 is the subclass which is most commonly used for experimental

work due to its abundance. It consists of twelve domains, four in each heavy chain and two in each light chain. The heavy chains have three constant and one variable domain labelled Constant-Heavy domain 1,2 and 3 and Variable-Heavy domain (C_H1, C_H2, C_H3 and V_H). In contrast, the light chains have only one of each: Constant-Light domain (C_L) and Variable-Light domain (V_L). Each domain, whether heavy or light, is between 110- 130 residues long, 12-13kDa in size and is folded into two twisted stacked beta sheets **(Figure 67, Figure 68)**. The sheets enclose an internal volume of hydrophobic residues in something which is referred to as an immunoglobulin fold (Schroeder Jr. and Cavacini 2010). One of these sheets has four anti-parallel strands whilst the other has three. They are connected via bends/loops with little secondary structure and stabilised by an internal disulphide bond. The residues involved in the sheets tend to be conserved whereas the residues in the loops are more diverse (Harris et al. 1996).

Each domain is bound to its neighbour via Van der Waals interactions. They all have close lateral sideways association except C_H2 which has a conserved sugar chain linked to it, weakening its bonds to its neighbours. C_H2 also contains the binding site for several important effector molecules *e.g.* Fc receptors and C1q from the complement system. These are protein complexes which bind to CH2 of IgG and actively induce the compliment pathway.

All of the constant domains are completely identical whereas the variable domains contain changeable antigen binding sites. There is an enormous range in diversity when it comes to these antigen binding sites, however, the two variable domains on each IgG molecule are identical making each antibody specific to one particular antigen. (Delves et al. 2006)**.**

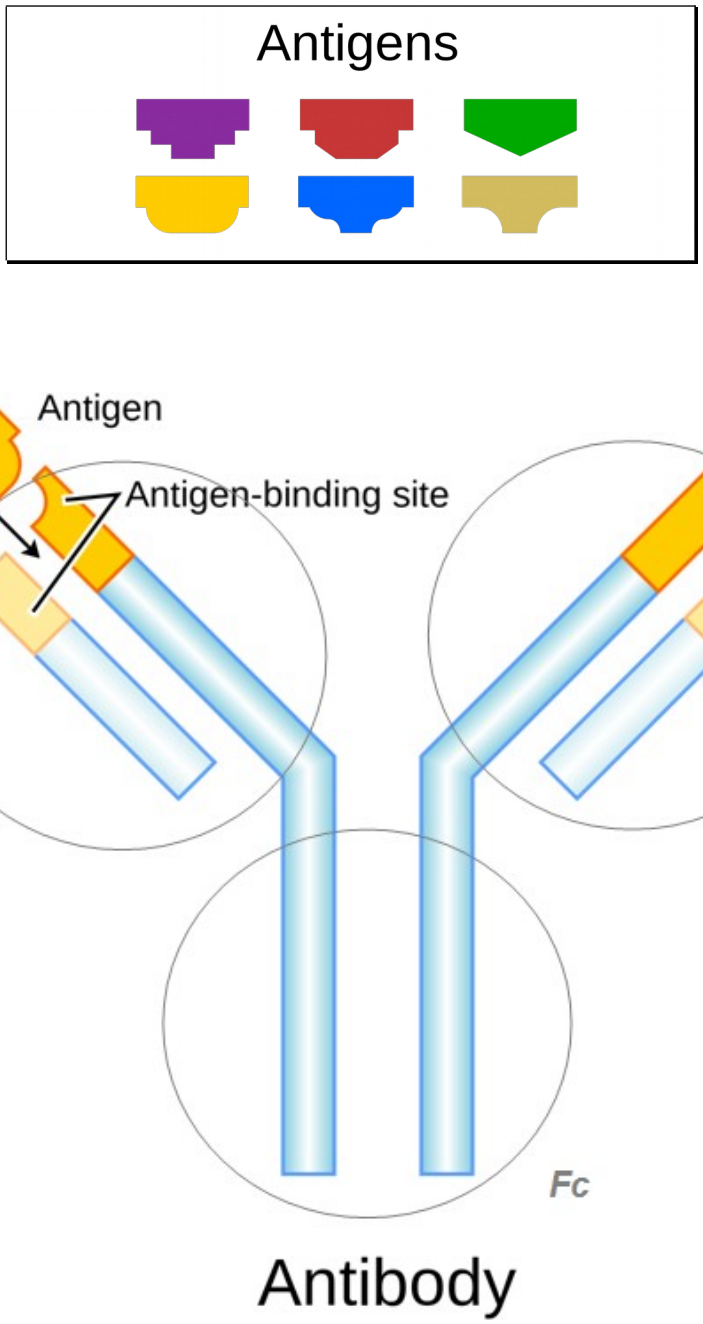


Figure 65: A simplified antibody structure with Fab and Fc regions labelled. The "Y" structure consists of four chains; two light and two heavy. The heavy chain spans both the Fab and FC region whereas the light chain lies exclusively in the Fab region. The Fab regions are identical and are responsible for antigen binding where as the remaining Fc unit is responsible for triggering the immune system through effector molecule binding (CK-12, 2017)

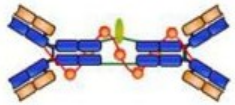
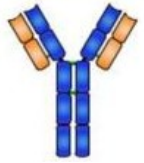
The five Immunoglobulin (Ig) subclasses					
	IgG	**IgA**	**IgM**	**IgD**	**IgE**
					
Heavy chain type	ɣ	α	μ	δ	ε
Molecular weight	150kDa	385kDa	900kDa	180kDa	200kDa
Number of antigen binding sites	2	4	10	2	2
Secretory component?	no	yes	no	no	no
J Chain	no	yes	yes	no	no
Crosses placenta	yes	no	no	no	no
Fixes complement	yes	no	yes	no	no
Fc binds to	phagocytes	Some polymorphonuclear leukocytes and lymphocytes	B cells	Not involved in compliment	Mast cells and basophils
Function	Main antibody of secondary immune responses, neutralizes toxins and opsinisation	Secreted into mucus, tears and saliva	Main antibody of primary response. Best at fixing complement	B cell receptor	Antibody of allergy and antiparasitic activity

Table 6: A table comparing the five antibody subclasses.

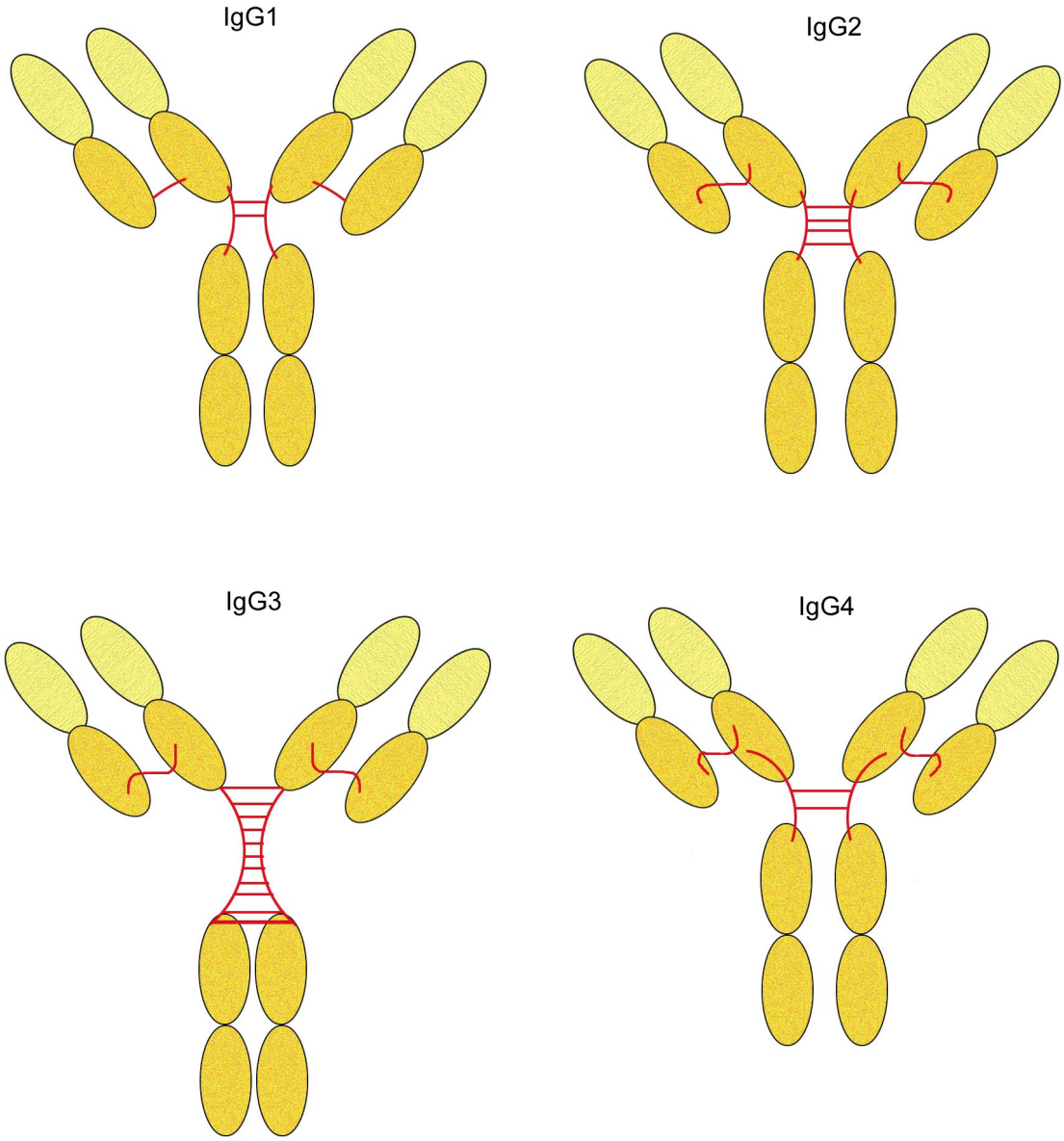


Figure 66: A comparison of the four IgG subclasses. Each are identical in all structural aspects expect for their hinge region. IgG2's hinge region consists of 4 inter-heavy chain disulphide bonds where as IgG3 consists of 11. Both IgG1 and IgG4 hinge regions hold two inter-heavy chain disulphide bonds, however, IgG1 is constructed of 15 amino acids where as IgG4 is constructed of 12 (Immunoglobulin G Deficiency, 2017).

	IgG1		IgG2		IgG3		IgG4	
General								
Molecular mass (kD)	146		146		170		146	
Amino acids in hinge region	15		12		62[a]		12	
Inter-heavy chain disulfide bonds	2		4[b]		11[a]		2	
Mean adult serum level (g/l)	6.98		3.8		0.51		0.56	
Relative abundance (%)	60		32		4		4	
Half-life (days)	21		21		7/~21[a]		21	
Placental transfer	++++		++		++/++++[a]		+++	
Antibody response to:								
Proteins	++		+/–		++		++[e]	
Polysaccharides	+		+++		+/–		+/–	
Allergens	+		(–)		(–)		++	
Complement activation								
C1q binding	++		+		+++		–	
Fc receptors								
FcγRI	+++[c]	65[d]	–	–	++++	61	++	34
$Fc\gamma RIIa_{H131}$	+++	5.2	++	0.45	++++	0.89	++	0.17
$Fc\gamma RIIa_{R131}$	+++	3.5	+	0.10	++++	0.91	++	0.21
FcγRIIb/c	+	0.12	–	0.02	++	0.17	+	0.20
$Fc\gamma RIIIa_{F158}$	++	1.2	–	0.03	++++	7.7	–	0.20
$Fc\gamma RIIIa_{V158}$	+++	2.0	+	0.07	++++	9.8	++	0.25
FcγRIIIb	+++	0.2	–	–	++++	1.1	–	–
FcRn (at pH < 6.5)	+++		+++		++/+++[a]		+++	

[a] *Depends on allotype.*

[b] *For A/A isomer.*

[c] *Multivalent binding to transfected cells. Adapted from Bruhns et al. (2).*

[d] *Association constant ($\times 10^6$ M^{-1}) for monovalent binding (2).*

[e] *After repeated encounters with protein antigens, often allergens.*

Table 7: A table comparing the properties of the four IgG subclasses (Vidarsson, Dekkers, and Rispens 2014).

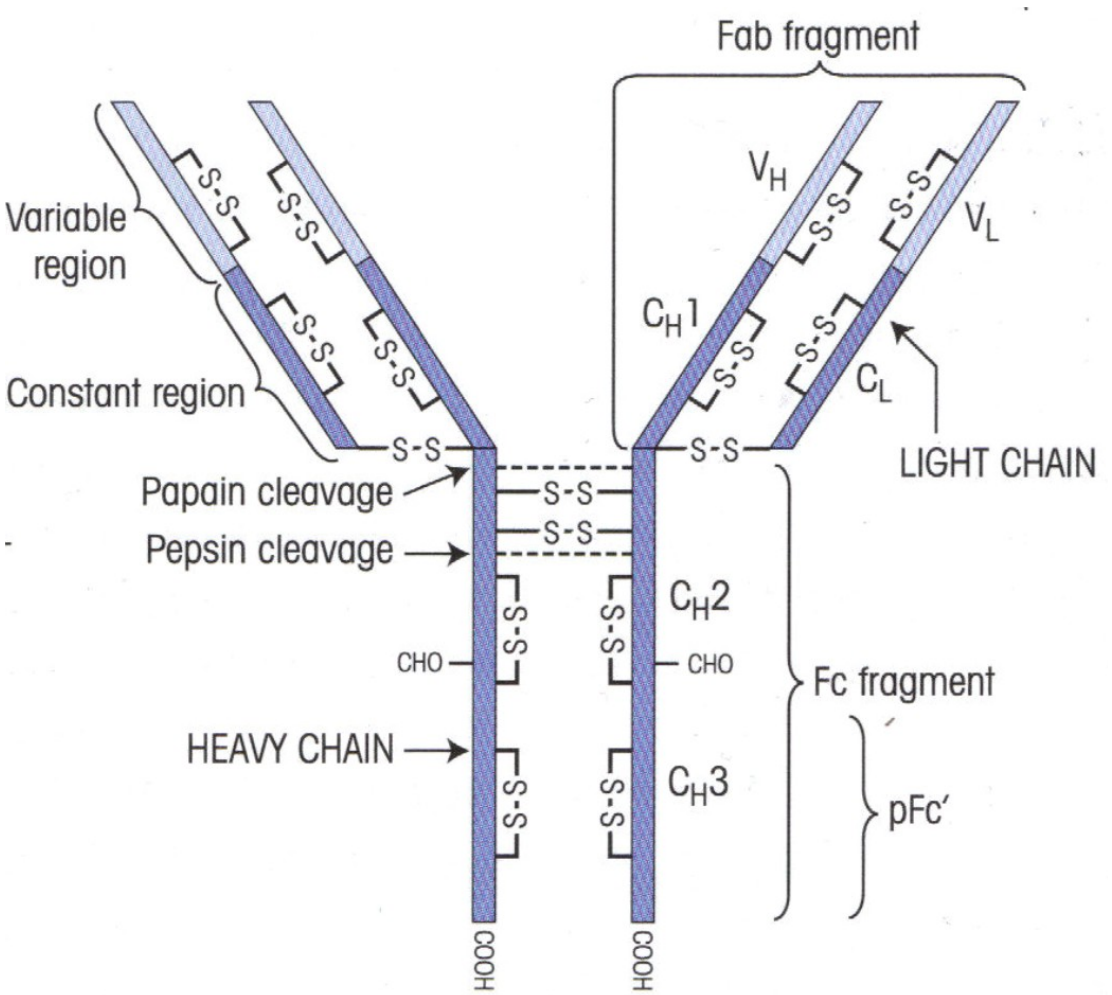


Figure 67: A simplified diagram of an IgG1 antibody. The 12 domains are divided into 2 light (L) chains and 2 heavy (H) chains connected by intermolecular disulphide bridges. The constant (C) and variable (V) domains are labelled as well as the enzyme cleavage sites of papain and pepsin (Delves et al., 2006).

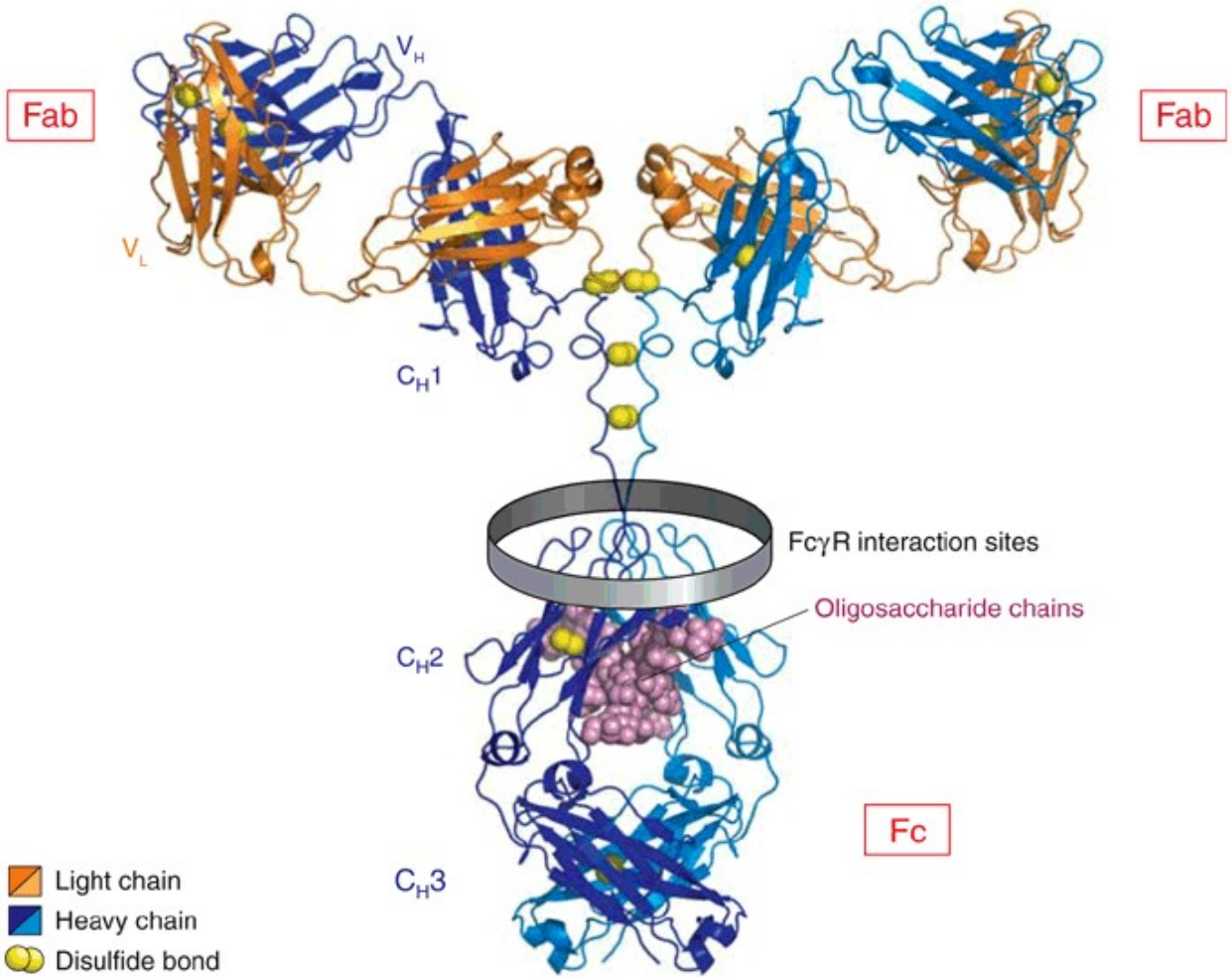


Figure 68: A ribbon diagram of IgG1. The heavy chains are represented in blue (dark ad light) and the light chains in orange. The yellow spheres represent the disulphide bonds which connect the chains to form the typical "Y" structure. The Fab and Fc regions are labelled alongside the oligosaccharides and immunoglobulin domains (Jefferis 2006).

VDJ Recombination

In order for an interaction to occur between antibody and antigen they must compliment each other in terms of shape and polarity. The antigen binding region of an antibody, also known as the hypervariable region, is made up of six Compliment Determining Region (CDR) loops; three from the V_H domain and three from the C_H domain **(Figure 69)**. Although made from a small genome (30k genes), the antibody is able to create enough variation in these loops for it to recognise millions of antigens. This is achieved through something called somatic VDJ recombination. This is where DNA is not inherited from the germ line cells but instead rearranges itself in somatic – in the cells of the body. Specifically it occurs during b-cell development in the bone marrow and involves the splicing of gene segments; mostly introns and some exons.

The three loops from the V_L domain are named L1,2 and 3 and the three loops from the V_H domain are named H1,2, and 3. The loops L2, H2 L3 and H3 are Beta-hairpins. These specific loops are responsible for linking successive anti-parallel strands of a single beta sheet whereas L1 and H1 form bridges from a strand in one of the beta sheets to a strand in the other. Five of the six CDR loops have a conserved suit of residues whose function is to maintain loop conformation in this hypervariable area. H3 is the loop with more variability in length, sequence and structure. H3 appears in the centre of the antigen binding site and makes significant interactions with its surrounding loops – *i.e.* its conformation depends on its environment (Lesk 2010) (Wood 2004).

The variable light and heavy chain loci of humans contain multiple gene segments which are joined through somatic recombination to produce the final variable region exon. The heavy chain variable region is constructed by the joining of three gene segments named variable (V), diversity (D) and joining (J). The light chain gene is constructed of only V and J and has options for both kappa and lambda variations. Diversity is increased further

through frequent mutations of the CDRs during B-cell maturation and from random pairing of heavy and light chains (Paul 2008).

V_H genes are mapped in chromosome 14 although some orphan IgH have been found in chromosome 15 and 16. The heavy chain has over 100 V genes, 27 D genes, 11 C genes and 9J gene giving a possibility of 24,000 conformations. The light chain lambda locus is found on chromosome 22 with 70 functional V genes and 7 functional J genes giving a possibility of 490 conformations. The remaining light chain kappa locus is located on chromosome 2 with 70V and 5J (350 combinations). The kappa locus, however, contains a large duplication of most the V genes (Liljas, A. 2009). Diversity is further induced by random combination of heavy and light chains. This allows a total of 24000 $^{(490 \times 350)}$ possible combinations. Additional point mutations and insertion/deletion of gene segments increases this further and provides protection against millions of antigens.

	Light lambda	Light kappa	Heavy
V	70+	70+	100+
J	7	5	9
C	7	1	11
D	-	-	27

A table demonstrating the genetic availabilities for VDJ recombination.

VDJ recombination is a regulated and ordered event. In the light chain it involves the DNA exon being constructed from a single V to J gene segment join. In the heavy chain, D is first joined to J and then V is combined to the DJ sequence. The re-arranged DNA is then transcribed, spliced and translated as normal (Delves et al. 2006) **(Figure 70)**.

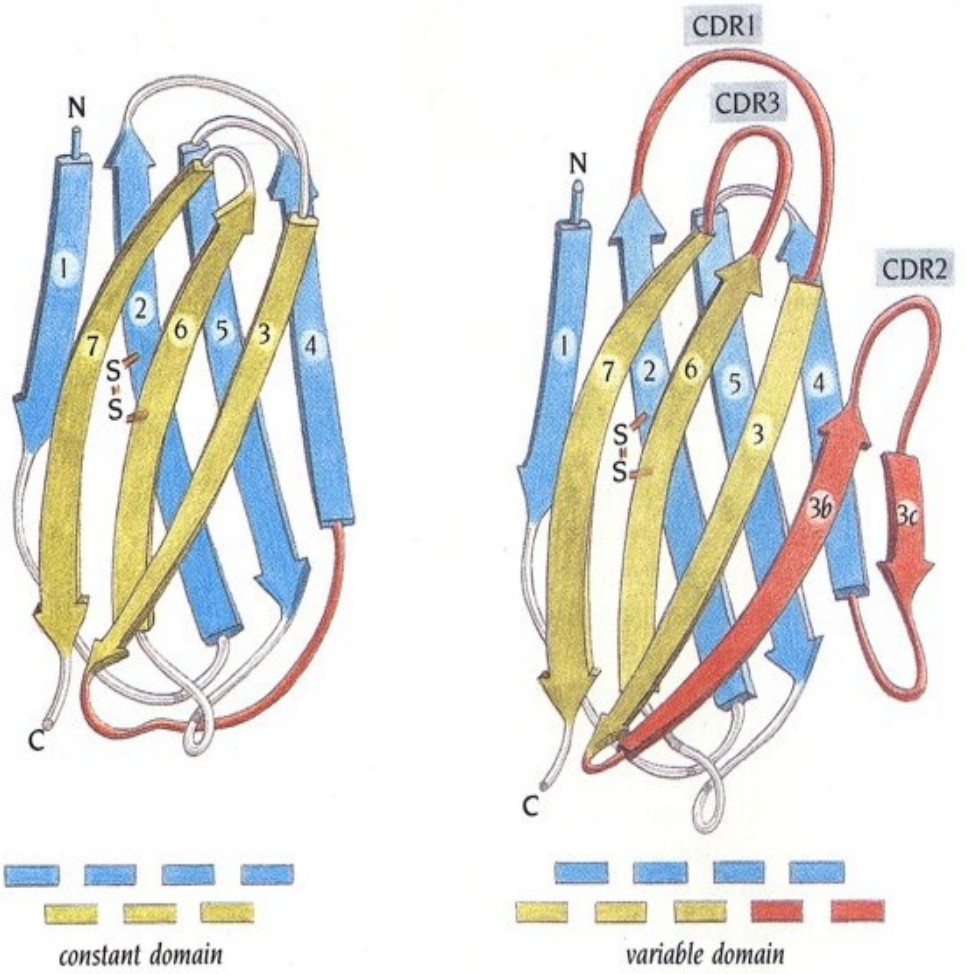


Figure 69: A ribbon diagram demonstrating the structure and packaging of the constant and variable immunoglobulin domains. Each domain consists of two twisted stacked beta sheets (blue and gold) enclosing an internal volume of hydrophobic residues. The discrepancies between the two domains are highlighted in red. The variable domain consists of two extra β-strands and three hypervariable CDR regions at the antigen binding site to enable higher levels of diversity. Regardless of being variable or constant, all domains require intra-molecular di-sulphide bonds for domain stability (labelled –S=S-) (Branden and Tooze 1991).

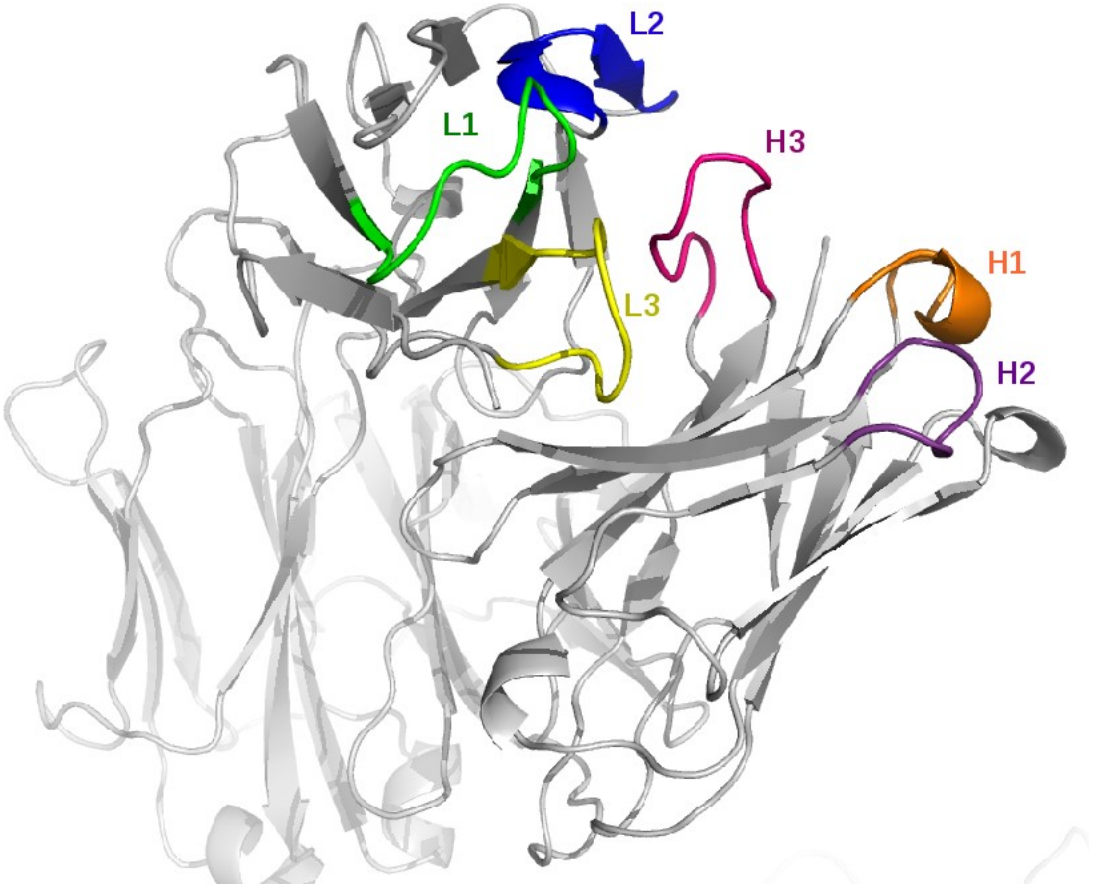
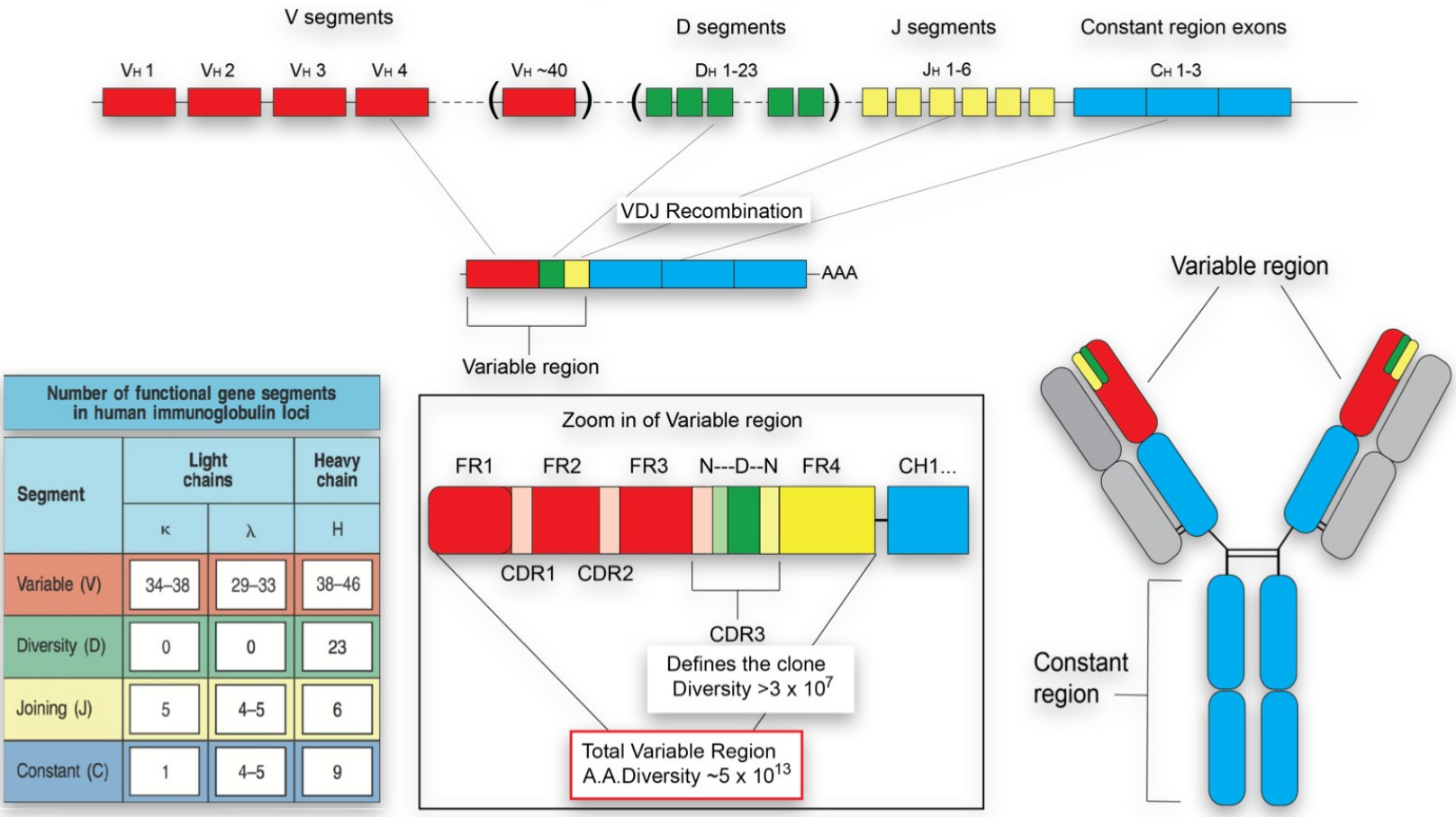


Figure 70: A ribbon diagram illustrating the CDR loops present at an antigen binding site of an antibody. Three loops are provided by the heavy chain (labelled H1,2 and 3) and three are provided by the light chain (labelled L1,2,and 3) giving a total of six hypervariable loops per arm. With each IgG1 molecule having two arms this gives an overall of 12 loops per IgG1 molecule (PDB code: 4NUJ).



A diagram illustrating VDJ recombination. In the light chain, the exon is constructed from a single V to J gene segment join. In the heavy chain, D is first joined to J and then V is combined to the DJ sequence. The re-arranged DNA is transcribed, spliced and translated as normal (Systems Immunology, 2017).

Antibody Preparations for X-Ray Crystallography

The anti-SAP antibody was provided by GlaxoSmithKlein in the form of an IgG1 molecule. In Due to the flexible nature of the IgG molecule, it was important that it be subject to fragmentation before any form of x-ray analysis could occur.

Antibody fragmentation can be achieved through the use of any protease; an enzyme that performs proteolysis by hydrolysing peptide bonds. However, for the purposes of this particular experiment, it necessary to cleave the IgG molecule specifically in its hinge region. Cleavage of this region results in rigid Fab and Fc production which is ideal for crystal growth.

The enzyme papain was chosen for IgG fragmentation. Papain, also known as papaya proteinase I, is a cysteine protease found in papaya whose cleavage site lies within the flexible hinge region of the IgG antibody. It is 212 residues long and is formed of two domains which are stabilised by three disulphide bridges. It is the cleft between these two domains which contains the active site. Like all cysteine proteases, papain cleaves proteins via hydrolysis of the peptide bond using a nucleophilic cysteine thiol in a catalytic triad mechanism (Kamphuis et al. 1986).

The first step in this reaction is the deprotonation of a cysteine thiol (cysteine 25) in the enzymes active site by the adjacent histidine (histidine 159). The deprotonated thiol attacks IgG's heavy chain between residues 224 and 225 (histidine and threonine respectively) via nucleophilic attack on the carbonyl. This causes cleavage to occur between histidine 224 and threonine 225 and as a result of this, a fragment of the substrate peptide is released (threonine 225 – lysine 330). Subsequently, histidine 224 becomes the new carboxyl terminus. It binds the cysteine thiol from the proteinase creating a thioester intermediate and causing the enzyme histidine to be restored to its original deprotonated state. Finally, hydrolysis occurs on the thioester creating a carboxylic acid moiety on the remaining fragment and regenerating the

enzyme. This process is then repeated between residues 233 (glutamic acid) and 234 (Leucine) (Lowe 1970).

It should be noted that the catalytic activity of papain can be inhibited in the presence of metal ions and certain organic compounds such as Imidazole (Sathish, Kaul, and Prakash 2000). Studies have shown papain to lose 50% of its activity in $2x10^{-4}$ M of cadmium ions and $4x10^{-4}$ M of zinc ions. At concentrations $1x10^{-3}$ M or higher of either metal or compound papain has shown to be completely inactivated. The loss in activity can be explained by the fact that all three have higher affinity for the active site than the substrate. Binding of these to the papain active site leads to small changes in the beta-structure content and causes a decrease in the thermal stability of the protein.

A protocol for papain digestion was found in a journal by Andrew and Titus in 2001. Here they carry out a 1:100 mass digestion for a total of 3 hours in order to obtain Fab and Fc fragments. These are later separated through the use of a HiTrap protein-A affinity column.

Protein-A is another molecule often used for single step, affinity chromatography. Protein-A is a 42kDa recombinant protein made in bacteria (*Staphylococcus aureus*). It is composed of 5 homologous immunoglobulin binding sites which fold into a 3 helix bundle. It has the ability to bind all human immunoglobulins through their Fc region by means of N-hydroxysuccinimide (NHS) activation; a process which causes the coupling of the IgG C-terminus carboxyl group with the amine group of protein-A via ester intermediates. This means that when subject to protein-A columns, the digested Fc fragments and the undigested IgG fragments all exhibit column binding whereas the remaining Fab fragments are eluted. In order to then remove the bound Fc and IgG fragments from the column, pH 4 buffer can be used. This low pH increases the hydrogen concentration in the column and removes the negative charge from protein A thus releasing the immunoglobulin carboxyl group.

Once separated from the other digestion components, the Fab fragments can be combined with the SAP so the make the complex. It is essential that the Fab fragments be added in excess to ensure full saturation of the SAP molecules therefore the molar masses of each component must be take into account. Fab fragments have a molecular mass of 50kDa where as the SAP pentamer has a molecular mass of 125kDa. This means that the Fab fragments are already in 2.5x molar excess. Therefore a minimum volume ratio of 2:1 SAP:Fab respectively would be necessary to ensure uniformity.

Upon forming of the complex, it can be separated from the excess Fab using High Performance Liquid Chromatography (HPLC). This particular form of chromatography separates its constituents via their size. A pumping system drives the mobile phase (the sample) through the HPLC column which contains a solid phase in the form of a porous resin. This resin allows smaller particles to pass through its complex network whilst excluding lager particles. In the case of these experiments, the larger SAP:Fab complex will have difficulty travelling through the central pore system of the column compared to the smaller excess Fab fragments. This will result in the complex travelling around the resin beads – a faster route – and being eluted first. Meanwhile, the excess Fab will be small enough to travel through the central aqueous system within the beads - a longer route - resulting in a later elution time.

Once separated and concentrated, the rigid complex will be able to undergo multiple rounds of crystallisation and optimisation in an attempt to form large, clear crystals suitable for x-ray analysis.

Methods

IgG Fragmentation Protocol

IgG Preparation

A sample of 103.5mg/ml anti-SAP IgG antibody was obtained from GlaxoSmithKlein (GSK). All impurities were removed from the sample via overnight dialysis in a cellulose membrane with a molecular weight cut off point of 100kDa. Dialysis was achieved at a ratio of 1:1000, IgG to sample buffer respectively (20mM sodium phosphate, pH 7).

In preparation of IgG fragmentation, an aliquot of dialysed IgG was diluted to 4ml of approximately 12mg/ml concentration using sample buffer containing EDTA (20mM sodium phosphate, 5mM EDTA, pH7). This was to ensure chelation of any metal ions present in the solution and thus prevent inactivation of papain further on in the protocol. The EDTA was subsequently removed via overnight dialysis against 500ml of sample buffer at 4°C and the prepared IgG was concentrated via centrifugation in an 18cm radius rotor Beckman Allegra X-12R centrifuge until 1ml of approximately 20mg/ml sample remained.

Papain Cleavage (1:100 Mass Digestion)

Immediately before digestion, powdered bovine papain was dissolved in digestion buffer (20mM sodium phosphate, 20mM Cysteine, pH 7) in order to produce a fresh enzyme stock 1mg/ml in concentration. An aliquot was taken and diluted to 0.2mg/ml using more digestion buffer. Of this, 0.66ml was added to the 1ml of prepared IgG to create a 1:100 mass ratio of papain to IgG respectively. The mixture was diluted to 4ml through further additions of digestion buffer and the mixture was left to incubate at 37°C for a total of

three hours. Samples of the digested mixture were subject to SDS PAGE analysis periodically on a 12% acrylamide gel so to check the outcome of the digestion. After completion of the incubation period, 30mM of iodoacetamide was added so to deactivate the papain and the digested sample was dialysed overnight against sample buffer in order to remove cysteines from the solution.

Papain Cleavage (1:100 Molar Digestion)

Immediately before digestion, powdered bovine papain was dissolved in digestion buffer (20mM sodium phosphate, 20mM Cysteine, pH 7) in order to produce a fresh enzyme stock 1mg/ml in concentration. An aliquot of this was taken and diluted to 0.2mg/ml using digestion buffer. Of this, a total of 0.17ml was added to 1ml of prepared IgG described above and this mixture was diluted to 4ml through the addition of more digestion buffer. The mixture was left to incubate at 37°C for a total of three hours. Samples of the digested mixture would be subject to SDS PAGE analysis periodically on a 12% acrylamide gel so to check the outcome of the digestion. Following completion of the incubation period, 30mM of iodoacetamide was added to the mixture so to deactivate the papain and the sample was dialysed overnight against sample buffer in order to remove cysteines from the solution.

Separation of Digestion Fragments

In order to purify the digested Fab fragments, a 1ml HiTrap protein-A affinity column was used. Following initial equilibration using sample buffer, the digested sample was passed through twice to ensure that all Fc and IgG fragments bound to the solid medium whilst the Fab fragments were eluted and collected. Following this, strength pH4 buffer (0.1M sodium citrate) was

used to detach and collect the Fc and IgG from the immobilised protein A. The column was then neutralised through the use of more sample buffer and stored in 20% ethanol.

Following separation, the purified Fab fragments were dialysed overnight against sample buffer at 4°C to remove phosphate from the solution - phosphate is known to precipitate in the presence of calcium - a component present in the SAP stocks. Following this, all digestion products were subject to SDS-PAGE analysis in order to ensure purity of each sample. Once purity was confirmed, the Fab samples from each digestion were combined and concentrated in an 18cm radius rotor Beckman Allegra X-12R centrifuge at 4°C.

Making the SAP:Fab Complex

At this point, it was unknown how many Fab fragments would bind to SAP but a maximum of five was designated per pentamer; one for each subunit. In order to ensure uniform complexes were to be formed, it was essential that the Fab fragments be added in excess. Fab fragments have a molecular mass of 50kDa where as the SAP pentamer has a molecular mass of 125kDa. This means that the Fab fragments are already in 2.5x molar excess. Therefore a minimum volume ratio of 2:1 SAP:Fab respectively would be necessary to ensure uniformity. It was chosen to add Fab in 4x molar excess as a precautionary measure to ensure full saturation of each SAP pentamer. Example below:

Example of making the SAP:Fab complex:

A total of 350µl Fab was collected (27.01mg/ml).

350/1000 x 27.01= 9.45mg of Fab was collected

Therefore, for a 4x molar ratio:
9.45 / 4 = 2.36mg of SAP stock (9.2mg/ml) would be needed

4.72 / 9.2 x 1000 = 257µl SAP to be added to Fab sample

In order to ensure the complex was in the pentameric form – the form representative of amyloid bound SAP - the designated SAP was incubated in calcium containing TC buffer for half an hour at room temperature before the Fab fragments were added. Subsequently, the mixture was gently mixed on a vortex and incubated for a further half an hour at room temperature before purification methods were carried out.

The complex was separated from the excess Fab using a Superdex 200 10/300 GL GE Healthcare Life Sciences HPLC column and the eluted complexes were concentrated via centrifugation in an 18cm radius rotor Beckman Allegra X-12R centrifuge at 4°C.

Crystallisation of the SAP:Fab Complex

Possible crystallisation “hits” for this new complex were determined using the robot crystallisation technique via the ‘mosquito’ programme. Crystallisation was performed using the method of vapour diffusion at room temperature in 1µl hanging drops at a ratio of 1:1 protein to well solution. A total of 96

different crystallisation solutions were trialled using the MD1-42 selection of the Molecular Dimensions commercial solutions.

The screens obtaining hits were replicated manually and optimised through changes in pH, temperature, precipitant and protein concentration. Crystals were harvested and flash-cooled in liquid nitrogen using 50% (v/v) glycerol as a cryoprotectant.

Data Collection and Processing

Diffraction patterns were obtained from a crystal grown in 0.2M NaCl, 0.1M potassium phosphate, 40% PEG 200 at a 1:1 ratio (protein:well solution) at a temperature of 4°C. Diffraction to approximately 4.4Å was observed at Diamond Light Source, Oxfordshire. A total of 450 Oscillation frames of 0.2° were measured with an exposure time of 0.08 seconds per frame (Beamsize 50x20μm, λ = 0.9763 Å). The data was processed automatically at Diamond.

Molecular Replacement

Matthews Co-efficient

Solvent calculations were performed using the 'Matthews - Cell Content Analysis' aspect of the CCP4i crystallography programme suite. Here, the cell parameters, space group and molecular weight of the complex were used to determine the solvent content. As the composition of the complex was not known for certain the molecular weight had to be estimated. It was decided to use the figure 375kDa as this is the total for an SAP molecule bound to a five Fab fragments.

Phaser

Initial phases were obtained by molecular replacement using the 'Phaser' aspect of the CCP4i programme suite. The search models used in the rotation and translation search of the 4.4Å data were a SAP pentamer from the refined structure of human SAP to 2Å with the calcium ions omitted (accession code 1sac; Emsley *et al.,* 1994) and Fab fragment refined structure of the human HIV-1 broadly neutralizing antibody PGT152 to 1.8Å (accession code 4NUJ; Blattner *et al.,* 2014).

Results

IgG Fragmentation

A sample of 103.5mg/ml anti-SAP IgG antibody was dialysed and digested using the protein papain at a mass ratio of 1:100 papain to IgG respectively. The mixture was left to incubate at 37°C for a total of three hours. The digestion was stopped via the addition of 30mM of iodoacetamide and the results of the digestion were presented on a 12% SDS acrylamide gel **(Figure 71)**.

The contents of each lane are as follows: 1) IgG control showing heavy chains at approximately 50kDa and light chains at 25kDa, 2) Digestion sample after 15 minutes, 3) Digestion sample after 30 minutes, 4) Digestion sample after 45 minutes, 5) Digestion sample after 60 minutes, 6) Digestion sample after 90 minutes, 7) Digestion sample after 120 minutes, 8) Digestion sample after 180 minutes, 9) BenchMark Pre-stained Protein Ladder.

The results of the gel clearly demonstrate that digestion of the heavy chain into Fab and Fc fragments occurs straight away. Lane 2, which represents the sample after only 15 minutes of digestion, clearly shows two additional bands representing the Fab and Fc fragments at approximately 25 and 30kDa respectively compared to the IgG sample in lane 1. The band 60kDa in size representing in-tact heavy chains disappears in lane 5 indicating that digestion is complete within one hour. When looking at the bottom area of the gel an issue becomes apparent. Over digested protein fragments are clearly visible in the form of smears in lanes 2-8. This suggests that over-fragmentation of the IgG molecule has occurred. This is an undesirable outcome as over-digestion can lead to low yields of Fab fragments. Due to the ‘trial and error’ nature of crystal making, high yields of Fab fragments are a necessity. They allow the preservation of the IgG stock and also allow for

more opportunities to attempt crystallisation if optimum conditions are difficult to find.

In order to resolve this issue of over-fragmentation, it was decided to attempt digestion using a 1:100 *molar* ratio instead (papain to IgG respectively). In order to perform a 1:100 molar digestion between the two components the molecular weights of each had to be taken into account. Papain has a molecular weight of 23kDa where as IgG has a molecular weight of 145kDa. This means that papain is already 6 times in molar excess. Therefore, in order to establish this particular molar ratio, the amount of papain had to be x600 less than that of IgG.

Upon digestion of IgG with papain at a molar ratio of 1:100 at 37°C for a total of three hours the following acrylamide gel was produced: **(Figure 72)**. The contents of each lane are as follows: 1) IgG control showing heavy chains at approximately 50kDa and light chains at 25kDa, 2) Digestion sample after 15 minutes, 3) Digestion sample after 30 minutes, 4) Digestion sample after 45 minutes, 5) Digestion sample after 60 minutes, 6) Digestion sample after 90 minutes, 7) Digestion sample after 120 minutes, 8) Digestion sample after 180 minutes, 9) BenchMark Pre-stained Protein Ladder.

Again, this gel clearly demonstrates that digestion occurs immediately. Bands representing Fab and Fc fragments are clearly seen at approximately 25 and 30kDa after only 15 minutes of incubation as represented by lane 2. The smear marks at the bottom of the gel are no longer present confirming that over-digestion is eliminated in the molar ratio. However, unlike in the 1:100 mass ratio, digestion does not actually reach completion. A clear band at approximately 60KDa can still be seen in lane 8 indicating in-tact heavy chains are still present at the end of the protocol.

Comparing these two gels it becomes apparent that the molar ratio provides a ‘cleaner’ digestion, however, this method does not merit in terms of time or yield. For a full digestion to occur, an estimated 6 hours of incubation would be needed. In contrast, the 1:100 mass digestion, although much quicker,

produces too much waste in the form of over fragmented IgG molecules. Both of these present a problem with respect to crystallisation. In order for SAP:Fab crystals to be obtained, multiple digestions need to be carried out in order to obtain sufficient stocks of Fab. This means the long digestion times of the molar protocol and the low yields of Fab fragments in the mass protocol are impractical.

In an attempt to solve these issues the digestion protocol was adjusted so to increase yields in a shorted time frame. Papain amounts were reduced to a third of their original 1:100 mass amounts and digestion time was reduced to one hour. Although this method would still result in the incomplete digestion of IgG and possibly the production of some over-digested fragments, these would not be to the extremes as seen in the previous digestions. Any over-digested fragments could be removed later during the purification process and left over IgG could be reused for subsequent digestions.

Using these newly adjusted digestion conditions, the digestion process was repeated a total of 9 times so to build up sufficient stocks of the Fab fragment. Following each digestion, the samples were submitted to SDS PAGE analysis to ensure successful digestion had occurred. An example is exhibited in **(Figure 73)**. The contents of each lane are as follows: 1) IgG control showing heavy chains at approximately 50kDa and light chains at 25kDa, 2) Digestion sample after 1 hour showing Fc and Fab fragments along with a faint band representing undigested heavy chains, 3) Purified Fab, 4) Purified Fc with faint lines representing undigested heavy chains, 5) BenchMark Pre-stained Protein Ladder.

The sample in lane 2 clearly demonstrates successful digestion of the IgG molecule. Strong bands at approximately 25kDa and 30kDa can be seen representing Fab and Fc respectively. A very faint band representing heavy chains can still be seen at approximately 60kDa however this is not seen in lane 3 indicating that the Fab fragments have successfully been separated from the IgG and Fc fragments through use of the protein-A affinity column.

The results of all 9 digestions are listed in the table below:

Digestion number	Initial IgG amount (mg)	Final Fab Amount (mg)	Final Fc amount (mg)	Yield (%)
1	7.5	3.5	0.7	56
2	20	3.6	3.1	34
3	14	6.1	3.7	70
4	22	9.0	7.3	74
5	23	12.0	5.8	77
6	20	5.5	6.9	62
7	20	9.8	6.2	80
8	18	6.0	6.4	61
9	21	6.7	3.1	46

Table 8: A table of anti-SAP IgG1 digestion yields. Yields were calculated through the combination of final Fab and Fc concentrations compared to the initial IgG concentration of that particular digestion. Component amounts were determined through use of a nano drop spectrophotometer set at a wavelength of 280nm. Units were converted using the Beer Lambert Law.

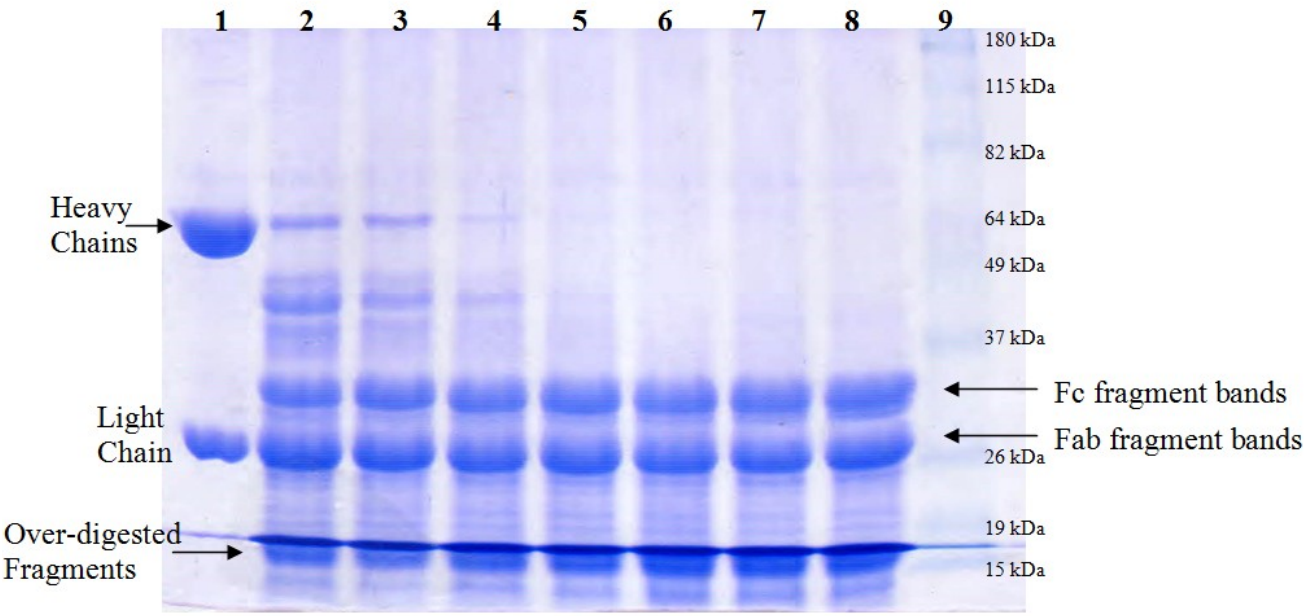


Figure 71: The Electrophoresis results of a 1:100 Mass Digestion - Papain to IgG respectively. A total of 10µl was added to each lane except in the case of the marker in lane 9 where only 5µl was added for preservation purposes. During preparation, samples were diluted to 1mg/ml using SDS containing laemmli buffer and boiled in order to enforce monomerisation and uptake of negative charge. Electrophoresis was carried out for 1 hour 20 minutes at 25 mA on a home made 12% acrylamide gel.

The contents of each lane are as follows: 1 – IgG control showing heavy chains at approximately 50kDa and light chains at 25kDa, 2 – Digestion sample after 15 minutes, 3 – Digestion sample after 30 minutes, 4 – Digestion sample after 45 minutes, 5 – Digestion sample after 60 minutes, 6 – Digestion sample after 90 minutes, 7 – Digestion sample after 120 minutes, 8 – Digestion sample after 180 minutes, 9 - BenchMark Pre-stained Protein Ladder.

Digestion of the heavy chain into Fab and Fc fragments occurs straight away as seen in lane 2 where clear bands representing each fragment are visible in the 15 minute sample. Also, the band representing in-tact heavy chain disappears in lane 5 indicating that digestion is complete within one hour. However, over digested protein fragments are visible at the bottom of the gel in the form of smears in lanes 2-8.

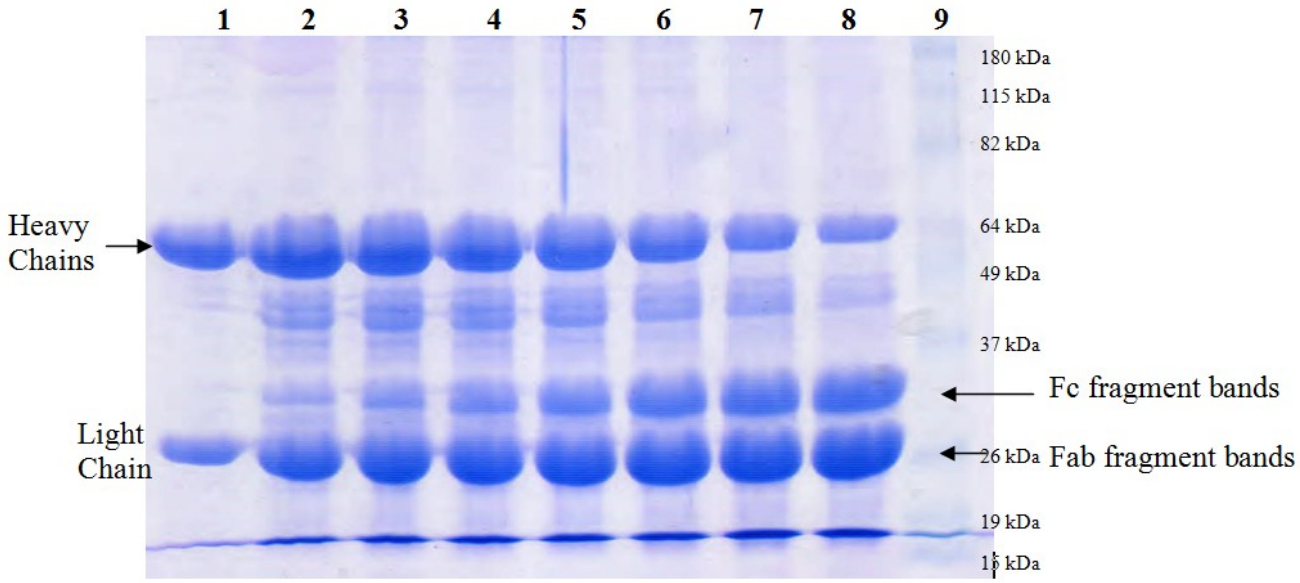


Figure 72: The Electrophoresis results of a 1:100 Molar Digestion - Papain to anti-SAP IgG respectively. A total of 10µl was added to each lane except in the case of the marker where only 5µl was added for preservation purposes. During preparation, samples were diluted to 1mg/ml using laemmli buffer containing sodium dodecyl sulfate (SDS) and boiled in order to enforce monomerisation and uptake of negative charge. Electrophoresis was carried out for 1 hour 20 minutes at 25 mA on a home made 12% acrylamide gel.

The contents of each lane are as follows: 1 – IgG control showing heavy chains at approximately 50kDa and light chains at 25kDa, 2 – Digestion sample after 15 minutes, 3 – Digestion sample after 30 minutes, 4 – Digestion sample after 45 minutes, 5 - Digestion sample after 60 minutes, 6 – Digestion sample after 90 minutes, 7 – Digestion sample after 120 minutes, 8 – Digestion sample after 180 minutes, 9 - BenchMark Pre-stained Protein Ladder (molecular weights of each band is labelled).

Digestion of the heavy chain into Fab and Fc fragments did not reach completion – a band representing heavy chains at approximately 60kDa is still present in lane 8 after 3 hours of digestion. However, digestion does occur straight away in lane 2 after only 15 minutes of incubation with no noticeable smear marks occurring at the bottom of the gel.

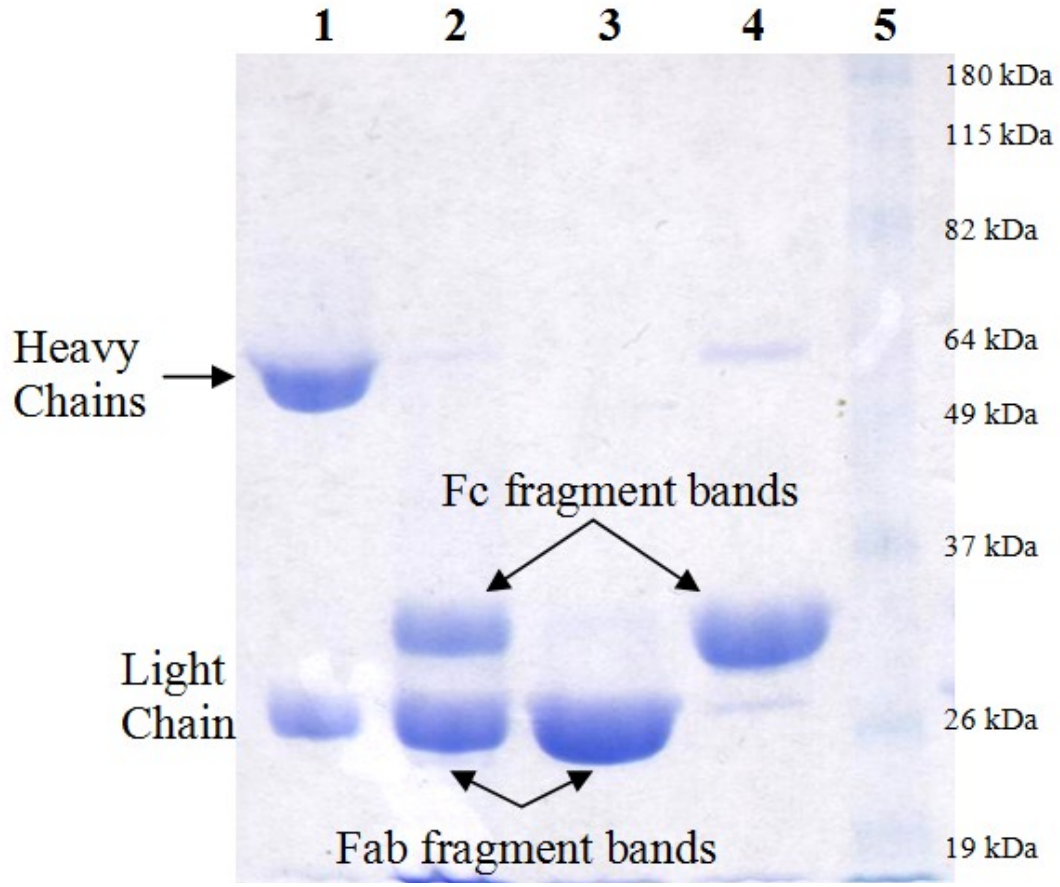


Figure 73: Typical electrophoresis results of an optimised anti-SAP IgG digestion. A total of 7.5µl was added to each lane except in the case of the marker where only 5µl was added for preservation purposes. During preparation, samples were diluted to 1mg/ml using sodium dodecyl sulfate (SDS) and boiled in order to enforce monomerisation and uptake of negative charge. Electrophoresis was carried out for 1 hour at 25 mA on a home made 12% acrylamide gel.

The contents of each lane is as follows: 1 – IgG control showing heavy chains at approximately 50kDa and light chains at 25kDa, 2 – Digestion sample after 1 hour showing Fc and Fab fragments along with a faint band representing undigested heavy chains, 3 – Purified Fab, 4 – Purified Fc with faint lines representing undigested heavy chains, 5 - BenchMark Pre-stained Protein Ladder (molecular weight of each band is labelled).

Making the SAP:Fab Complex

The purified samples of Fab fragments were combined and concentrated on an 18cm radius rotor Beckman Allegra X-12R centrifuge at 4°C until a 350µl sample 27.1mg/ml Fab was produced. This was gently mixed with 257µl SAP to ensure the Fab fragments were in x4 molar excess and incubated for half an hour at room temperature to ensure binding.

High Performance Liquid Chromatography (HPLC) was used to separate the complex from the excess Fab fragments. Two distinct peaks can be seen on the elution graph. The larger first peak represents the pure SAP:Fab complex and the second smaller peak represents the excess Fab fragments which did not bind.

Crystallisation of the SAP:Fab Complex

Possible crystallisation "hits" for this new complex were determined using the robot crystallisation technique via the 'mosquito' programme. Crystallisation was performed using the method of vapour diffusion at room temperature in 1µl hanging drops at a ratio of 1:1 protein to well solution. A total of 96 different crystallisation solutions were tested using the MD1-42 selection of the Molecular Dimensions commercial solutions.

Over nucleation of the SAP:Fab complex was experienced in the following six conditions after 2 weeks of incubation at room temperature:

A6 - 0.2M Sodium chloride, 0.1M sodium/potassium phosphate, pH 6.5, 25% (w/v) PEG 1000
C2 - 0.1M Sodium citrate, pH 4.5, 20% (w/v) PEG 4000
C7 - 0.1M Sodium citrate, pH 5.6, 20% (w/v) PEG 4000, 20% (v/v) 2-propanol
D7 - 0.1M Sodium citrate, pH 5.5, 15% (w/v) PEG 6000
E2 - 0.1M Sodium citrate, pH 5.0, 8% (w/v) PEG 8000

E11 - 0.1M Sodium citrate, pH 5.0, 20% (w/v) PEG 8000

These “hits” were replicated manually using the method of vapour diffusion at room temperature using 4µl hanging drops at a ratio of 1:1 protein to well solution respectively. Each particular ‘hit’ was repeated exactly as per the commercial solution in a single well of its own individual crystallisation tray. The remaining 15 wells of each tray, although similar to the ’hit condition’ were subject to slight alterations in pH, precipitant concentration and protein concentration as a way to decipher preferred growth conditions.

Vast amounts of over-nucleation were noticed in these manual trays after two weeks of incubation. The pattern of nucleation was observed and from this, preferred well conditions were estimated and created on new crystallisation trays. Due to the large amounts of over nucleation, it was decided to submit the trays to colder temperatures of 4°C so to encourage slower growth and thus encourage crystal formation.

Following eight rounds of optimisation and more than a year of trial and testing, crystals were finally produced in two of the six crystal screens. Those from Crystal screen #1 were produced via the hanging drop method at 4°C at pH’s of 4.5 and 5.5 , 0.1 sodium citrate, 13-14% PEG 3000 at ratios of 1:1 and 2:3 protein to well solution respectively. The crystals were of needle morphology and initially formed following 3 weeks of incubation. A further month was needed in order to allow crystal growth **(Figure 74,)**. Those from Crystal screen #2 were produced via the hanging drop method at 4°C at pH’s of 6 and 8, 0.2M sodium chloride, 0.1M potassium phosphate, 30-45% PEG 200 at ratios of 1:1 and 2:3 protein to well solution respectively. The crystals were of cube morphology however they contained cracks and were very sensitive to changes in temperature and environment. They formed following 4 weeks of incubation **(Figure 75,)**.

Although crystals were produced in both trays, the crystals from screen 1 were deemed too thin for x-ray analysis so were not flash frozen. In order to provide safe freezing conditions for the remaining cube crystals, various

cryoprotectants from the Molecular Dimensions CryoProtX™MD1-61 kit were trialled. The majority of these, unfortunately, resulted in extreme cracking of the crystals and many were destroyed. Ultimately a simple solution of 50% glycerol was found to be the best suited for this crystal type. It should be noted, however, that upon transferring the crystals to this solution, further cracking could still be exhibited in the lattices if the change of environmental conditions occurred too fast.

Crystal screen #1

Protein: *SAP:Fab complex*
Conditions: *0.1M Na Citrate, PEG 3K*
Ligand: *Calcium, PE*

	pH 4.5			pH 5.5			
	1	**2**	**3**	**4**	**5**	**6**	PEG 3K
A							12%
B			o/n		*HIT		13%
C	*HIT	o/n	o/n	o/n	*HIT	o/n	14%
D							15%
Drop ratio Protein: well	2:2	2:3	2:4	2:2	2:3	2:4	

Figure 74: A representation of a manually prepared crystal screen. This particular screen illustrates the optimised conditions for the SAP:Fab complex grown in sodium citrate and PEG3000 over a pH range of 4.5-5.5. (o/n = over nucleation, *HIT = crystal(s)).

Crystal Screen #2

Protein: *SAP:Fab complex*
Conditions: *0.2M NaCl, 0.1M Potassium phosphate, PEG 200*
Ligand: *Calcium, PE*

	pH 6		pH 7		pH 8		
	1	**2**	**3**	**4**	**5**	**6**	PEG 20K
A	*HIT						30%
B	o/n					*HIT	35%
C	o/n				*HIT		40%
D		o/n					45%
Drop ratio Protein: well	2:2	2:3	2:2	2:3	2:2	2:3	

Figure 75: A representation of a manually prepared crystal screen. This particular screen illustrates the optimised conditions for the SAP:Fab complex grown in sodium chloride, potassium phosphate and PEG 200 over a pH range of 6-8. (o/n = over nucleation, *HIT = crystal(s)).

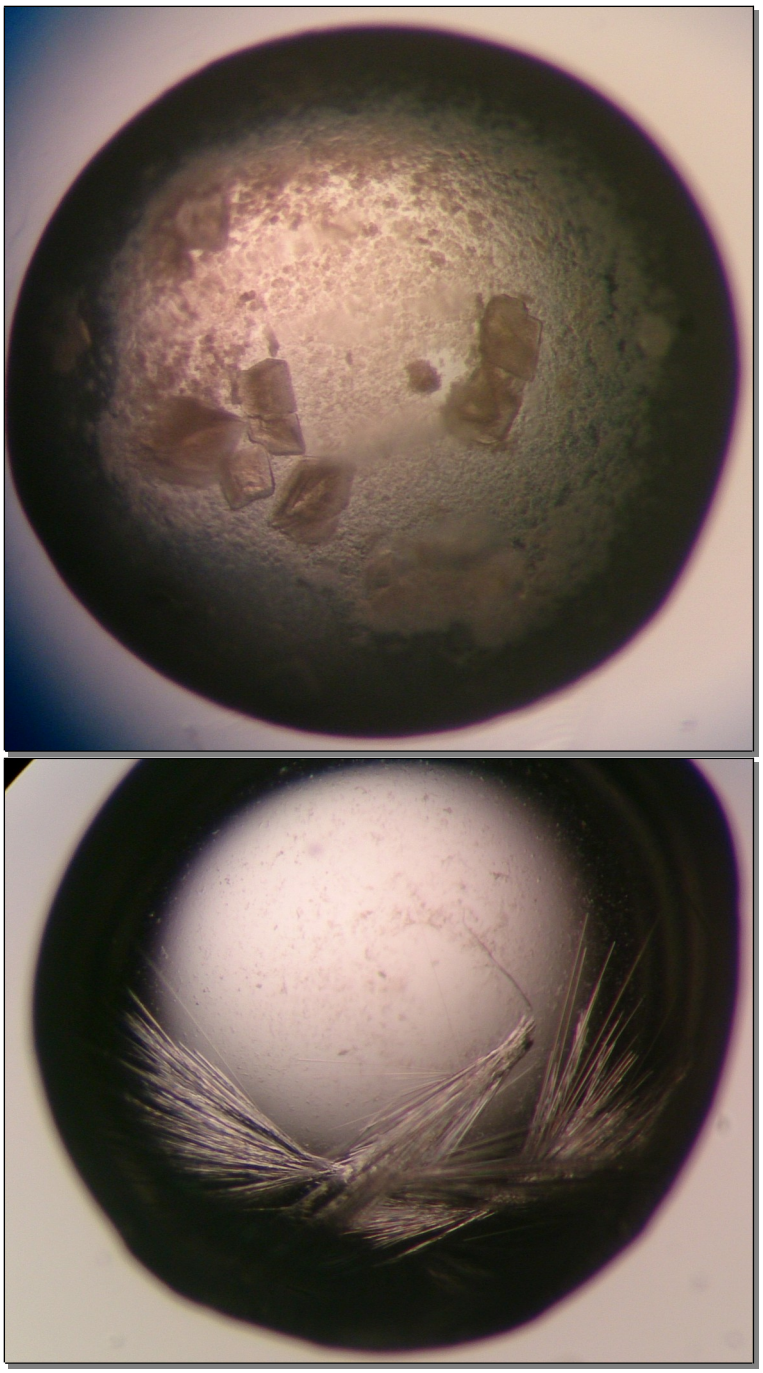


Figure 76: Comparisons of SAP:Fab crystals grown in different well solutions. Above – Example of crystals grown in 0.2M NaCl, 0.1M potassium phosphate, 40% PEG 200 at a 1:1 ratio (protein:well solution). Well 'C5' from Crystal Screen 3 produced large square and rectangular crystals approximately 150 x 150µm in size. Below – Example of crystals grown in 0.1M Na Citrate, 14% PEG 3K at a ratio of 2:3 (protein:well solution). Well 'C5' from Crystal Screen 1 produced long, needle-like crystals approximately 5 x 750µm in size.

Data Collection and Processing

Of the14 cube shaped crystals from screen 2, only two survived the flash freezing process and were taken to Diamond Light Source for diffraction. These crystals were both approximately 150 x 150µm in size and grown in a solution containing 0.2M sodium chloride, 0.1M Potassium phosphate, 35% PEG 200.

Upon arrival to Diamond it was discovered that one of the crystals had not survived transportation, however, diffraction patterns were obtained for the remaining crystal. Diffraction data was collected to 4.4Å (Beamsize 50x20µm, λ = 0.9763 Å). Oscillation frames of 0.2° were measured with an exposure time of 0.08 seconds per frame to produce a total of 450 diffraction images. The data was auto-processed by Diamond and the following unit cell dimensions were proposed: a = 176.22, b = 217.09, c = 259.93, α = 90.00, β = 90.00, ɣ = 90.00. This suggests that the crystal belonged to the Orthorhombic space group P21 ($P\ 2_1\ 2_1\ 2_1$).

Parameter	Value
Space group	P21
Unit cell	a = 176.22, b = 217.09, c = 259.93, α = 90.00, β = 90.00, ɣ = 90.00
Resolution (Å)	4.4
Measured reflections	450
Multiplicity	3.4
Completeness (%)	99.7
Mean (I)/sd(I) (outer resolution shell)	9.6
Solvent Content (%)	39.43

Data processing statistics for the SAP:Fab crystal

Molecular Replacement

Solvent calculations were performed using the 'Matthews coefficient – Cell Content Analysis' function of the CCP4i programme suite. Here, the cell parameters, space group and molecular weight of the complex were used to determine the solvent content. As the composition of the complex was not known for certain an estimated molecular weight was inputted. A molecular weight of 375kDa was chosen as this is the mass of a fully saturated SAP molecule, *i.e.* an SAP molecule with five Fab fragments bound.

The results of the Matthew's coefficient for an estimated molecular weight of 375kDa were as follows:

Nmol/asym	Matthews Coeff	%solvent	P(4.41)	P(tot)
1	6.63	81.46	0.00	0.00
2	3.31	62.91	0.24	0.16
3	2.21	44.37	0.75	0.83
4	1.66	25.83	0.00	0.00
5	1.33	7.29	0.00	0.00

A table showing the results of the Matthew's coefficient calculation of the SAP:Fab dataset.

On average protein crystals should have a solvent content of around 43%, which equates to a Matthews coefficient of between 1.7 and 3.5 cubic A/Da. According to the table above, the closest solution obtained was one of 44.37%. This solution has 3 molecules in an asymmetric unit and exhibits a Matthew's coefficient of 2.21.

Initial phases were obtained by molecular replacement using the 'Phaser' aspect of the CCP4i programme suite. The search models used in the rotation and translation search of the 4.4Å data were an SAP pentamer from the refined structure of human SAP to 2Å with the calcium ions omitted (accession code 1sac; Emsley *et al.,* 1994) and Fab fragment refined structure of the human HIV-1 broadly neutralizing antibody PGT152 to 1.8Å (accession code 4NUJ; Blattner *et al.,* 2014)

After 24 hours of calculating, Phaser confirmed that there were indeed 3 molecules in the asymmetric unit. The results of the rotation function showed that 25 solutions had been found:

Number	LLG	z-score
1	90.83	5.11
2	83.07	4.83
3	79.89	4.72
4	78.97	4.69
5	75.67	4.58
6	73.30	4.49
7	68.93	4.34
8	68.71	4.33
9	66.53	4.26
10	59.30	4.00
11	52.26	3.76
12	50.83	3.71
13	42.40	3.41
14	41.67	3.39
15	41.34	3.38
16	38.42	3.27
17	37.86	3.25
18	36.82	3.22
19	36.53	3.21
20	35.83	3.18
21	34.28	3.13
22	34.04	3.12
23	33.42	3.10
24	32.85	3.08
25	32.71	3.07

A table showing the 25 possible rotation solutions as calculated by the Phaser programme.

The strength of each solution is represented by the Log Likelihood Gain (LLG) and the ‘z-score’. The LLG is an indication of how much "better" the solution is compared to a random solution where as the z-score shows you how many deviations the solution is above the mean. In both cases a higher value indicates a good solution. An LLG of 0 or less means that your solution is no better or even worse than a random selection of atoms where as a z-score of less than 5 indicates that a solution may not be possible. The best rotation solution as presented by Phaser had an LLG of 90.83 and a z-score of 5.11. This suggests that the solution may be solvable but this is not guaranteed.

The program took the top rotation function (SET 1) and found 9 possible translation solutions. The results are as follows:

Number	LLG	z-score
1	50.09	4.17
2	43.11	3.83
3	36.68	2.64
4	35.36	1.06
5	33.73	1.85
6	30.15	1.53
7	27.20	1.78
8	22.67	1.47
9	20.93	1.10

A table showing the 9 possible translation solutions as calculated by the Phaser programme.

Again the strength of each solution was represented by the Log Likelihood Gain (LLG) and the 'z-score'. The best solution as presented by Phaser had an LLG of 50.09 and a z-score of 4.17. This suggests that this particular data set is no better than a random solution of atoms and cannot be solved. Attempts to model the structure using the 'coot' function from the CCP4i programme suit were carried out in spite of this, however, it became clear very early on that refinement would indeed not be possible. The fit of the density angles into the proposed map demonstrated almost no validation whatsoever.

Discussion

The aims of this chapter were as follows:

- To undergo multiple rounds of anti-SAP antibody digestion so to obtain a stock of Fab fragments
- To Combine SAP with excess Fab fragments in order to create a fully saturated SAP:Fab complex
- Find optimal conditions to allow for crystal growth of said complex
- Submit the complex crystal to x-ray analysis in order to obtain a diffraction data set
- Solve the diffraction data set using the CCP4i programme suite so to decipher the interaction between the SAP and Fab fragments.

Of these five steps, four were successfully carried out and executed. Optimum conditions for SAP:Fab formation were deduced and a crystal was subsequently submitted to x-ray diffraction. Issues arose with regards to solving the structure, however, due to the low LLG and z-scores calculated from the translation function of the Phaser programme.

Building and refining a protein model is not an exact science. Rather, it is a subjective process, governed by experience, prejudices, expectations and local practices. This means that errors in this process are almost unavoidable, but it is the crystallographer's task to remove as many of these as possible. It is a difficult process due to the limited resolutions and imperfect phases often presented by the diffraction data.

With high resolution data and good phases, the resulting model is approximately 95% a consequence of the data; however, once the resolution becomes worse than 2 Å, this balance shifts and protein models can often be a result of the crystallographer’s imagination. In general, the clarity and interpretability of electron density maps depends on the resolution of the diffraction data even when based on accurate phases. This is because the

relative error of diffraction intensities in the resolution shell of 4.5-4.0 for crystals diffracting to 4 Å is much larger than for crystals diffracting to 1.5 or even 2 Å.

Resolution is defined as the minimum spacing (d) of crystal lattice planes that still provide measurable diffraction of x-rays. It measures the minimum distance between structural features that can be distinguished on an electron density map and thus defines the level of detail that can be obtained. The smaller the spacing the more reflections available to define the structure and thus the higher the resolution of the data. In general, data to a resolution of 1.99Å or higher is considered good. Those between 2 and 3 Å are generally considered medium whereas anything beyond 3 Å is typically classed as low resolution data (Wlodawer et al. 2008).

The data achieved from the SAP:Fab complex exhibited a resolution of 4.4Å. Although this is considered low in terms of model building, crystals of lower data have previously been solved. The most famous of these is the myoglobin structure which was presented at a resolution of approximately 6 Å in 1958 by Kendrew *et a*l. This was the very first crystal structure to ever be solved. Due to its low resolution, much of the atomic detail known today about myoglobin was missing from this data however a very rough idea on its structure was presented.

Similar ideas can be presented with regards to the SAP:Fab structure. The estimated images as presented in ‘coot’ give an idea as to how the two molecules interact. It would appear that SAP binds to a single Fab fragment rather than the 5 previously guessed. Of which, the Fab fragment binds to the edge of the SAP structure with a slight propensity towards the calcium binding B-face. The most plausible reasoning behind this low number of Fab fragments interactions can be explained by allosteric conformation changes. This is the process where upon binding of a Fab fragment to one of the SAP monomeric units, conformational changes occur on the rest of the SAP structure. These alterations in structure prevent any further interactions between SAP and Fab fragments. The precise residues involved in Fab

binding and the structural changes which occur, unfortunately, cannot be determined using this information **(Figure 77)**.

The downfall of the SAP:Fab structure determination is thought to lie within the experimental aspect of this chapter. One must remember that in order for successful diffraction to occur several stages need to be executed and optimised. The right environmental conditions must be fond to encourage large, clear crystal growth. Once grown, the crystals must be flash frozen with liquid nitrogen in order to resist radical damage from the x-ray beam. This must be done with a cryoprotectant which does not cause any damage to the crystal itself. If any one of these stages experience an issue then it is likely the finishing data will be compromised. The SAP:Fab crystal had demonstrated a cracked morphology within its tray conditions; a property which was only enhanced through the transference to a cryoprotectant medium. This is clearly the root of the data inadequacy.

The most obvious solution for this problem would be to optimise the crystal screens further and thus produce more crystals. These could be subject to further trial and error in an attempt to find alternate cryoprotectants which provide a safer environment for flash freezing and hopefully allow for better diffraction.

This solution was attempted. When replicated and subject to another round of optimisation, cube shaped crystals were, again, produced in solutions containing 0.2M sodium chloride, 0.1M Potassium phosphate and PEG 200. Unfortunately, these screens were knocked over and destroyed by a third party before freezing. Stock solutions, at this point had been depleted and another tray could not be produced in time.

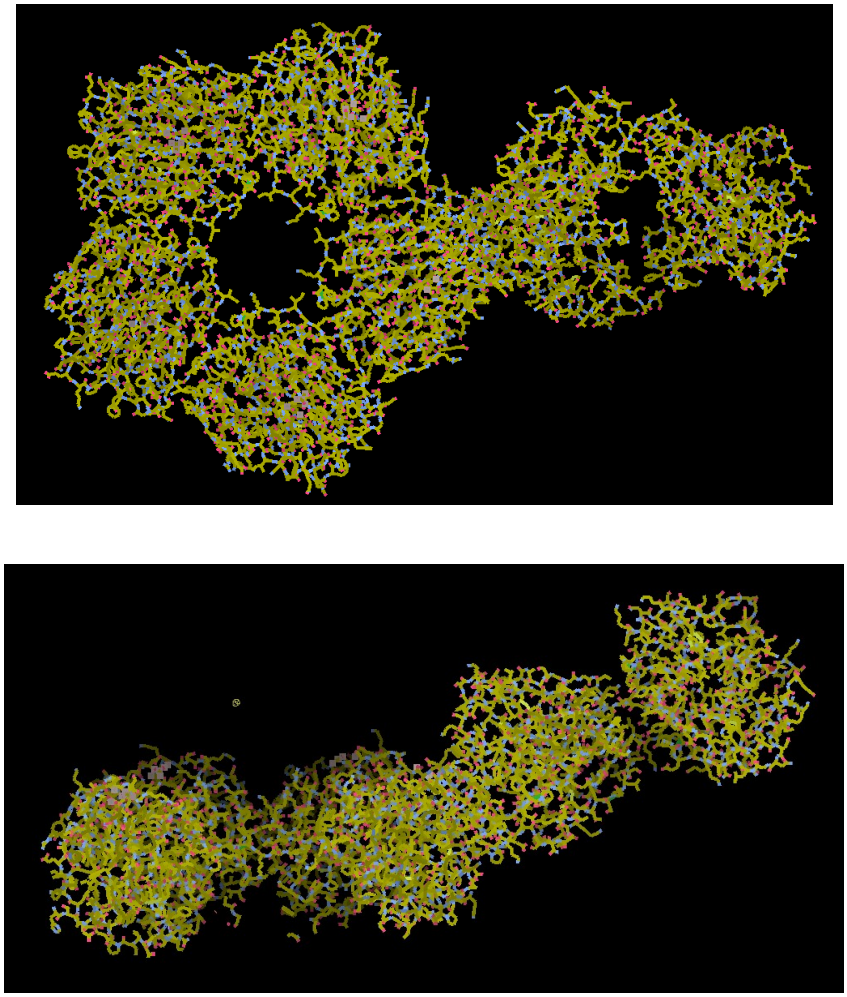


Figure 77: Images of the estimated SAP:Fab binding as demonstrated in the 'coot' function of the CCP4i programme suite. It is clearly demonstrated that a single Fab fragment binds to the pentameric SAP structure. (Above) The structure from a flat perspective of the B-face of SAP. (Below) The structure from a side perspective. Here you can see that the binding of the Fab fragment leans towards the B-face of the SAP molecule.

Final Remarks

The purpose of this thesis was to determine whether SAP exhibits *both* pro-fibrillogenic and anti-fibrillogenic activities as argued by Hamazaki and Janciauskiene. Using insulin fibres made at neutral pH this was proven to be true. ThT, Congo red, CD and TEM analysis all confirmed this. SAP pentamers enhance fibre growth whereas the decamers prohibit it. As of yet, this dual role has not been confirmed with any other amyloid prone protein due to the difficulties faced when trying to produce said proteins 'in house'. However, with a few more trials it would almost certainly be proven true. The morphologies of these new fibres would likely differ from that of insulin possibly causing SAP to have either a bigger or smaller impact on them.

By acting as a reducer of amyloid fibre mass, the possibility of using SAP decamers as a treatment for amyloid can be taken into consideration. Thus far, however, SAP has only demonstrated its enhancing capabilities *in vivo* and not its molecular chaperone behaviour. This fact, one could argue, makes the insulin results irrelevant. If SAP behaviour cannot be replicated *in vivo* this reducing role can be considered immaterial. That being said, if it could be proven that SAP has the *potential* to behave as a chaperone *in vivo,* then attempts could be made to facilitate this and thus a treatment against amyloid related disease could be developed. Not only is SAP universal to all forms of amyloid - therefore has the capability to act against every form of amyloid- , it is also a component already present in the human body so is unlikely to impose further damage to the host.

By using the decamers to reduce fibre mass one would regain tissue function and reduce symptoms of the patient. It may be true that the decamer abilities are limited to reducing amyloid fibres and not eradicating them, nevertheless, the lowering of deposit masses is enough to prolong the lives of these patients and reverse their symptoms. Remember, amyloid is not cytotoxic and in low concentrations amyloid is not life threatening. For that reason,

prolonged administration of an SAP treatment could be used until correct diagnosis has occurred and the relevant treatment obtained.

Of course there are factors which would need to be taken into consideration before such a treatment could be concocted. Firstly, if SAP decamers were able to form in the serum would they not already be present? Serum has a physiological pH of 7.35-7.45 with a low calcium concentration of 8.5-10.2mg/dL of which only 4.8-5.7mg/dL exists as free ionised ion in serum. This is not sufficient to saturate the SAP when their diverse molecular masses are taken into consideration and therefore one should expect SAP to exist as a decamer but it does not - Why? Is it simply a matter of pentamer binding with alternate serum proteins which prevent decamer formations or is it due some other unknown reason? Secondly, if SAP decamers were successfully administered into a patient, what would prevent them from reverting back to the pentamer formation and thus increase stability and growth rates of the existing fibres?

Although amyloid diseases are considered rare, they are likely more common than previously thought. The wide range of symptoms associated with amyloid disease often leads to misdiagnosis and thus there is a lack of literature regarding this category of diseases. The field has been grossly under researched as a result of this. So far 36 proteins have been associated with amyloid diseases and this number is continually increasing. Each of these diseases requires unique treatment to combat its unique set of symptoms and any delay can easily result in irreversible if not fatal consequences. Not enough is known about amyloid in general let alone each specific type to allow for effective, individual treatment plans. Because of this, the idea of having a general SAP treatment plan is extremely appealing.

References

Adamski-Werner, Sara L., Satheesh K. Palaninathan, James C. Sacchettini, and Jeffery W. Kelly. 2004. “Diflunisal Analogues Stabilize the Native State of Transthyretin. Potent Inhibition of Amyloidogenesis.” *Journal of Medicinal Chemistry* 47 (2):355–74. https://doi.org/10.1021/jm030347n.

Al, Liljas Anders Et. 2009. *Textbook Of Structural Biology*. Pap/Cdr edition. Hackensack, NJ: World Scientific Publishing.

Alun W Ashton, Mark K. Boehm. 1997. “Pentameric and Decameric Structures in Solution of Serum Amyloid P Component by X-Ray and Neutron Scattering and Molecular Modelling Analyses.” *Journal of Molecular Biology* 272 (3):408–22. https://doi.org/10.1006/jmbi.1997.1271.

Ancsin, John B. 2003. “Amyloidogenesis: Historical and Modern Observations Point to Heparan Sulfate Proteoglycans as a Major Culprit.” *Amyloid : The International Journal of Experimental and Clinical Investigation : The Official Journal of the International Society of Amyloidosis* 10 (2):67–79. https://doi.org/10.3109/13506120309041728.

Andersen, Dana C, and Lynne Krummen. 2002. “Recombinant Protein Expression for Therapeutic Applications.” *Current Opinion in Biotechnology* 13 (2):117–23. https://doi.org/10.1016/S0958-1669(02)00300-2.

Ando, Yukio, Teresa Coelho, John L. Berk, Márcia Waddington Cruz, Bo-Göran Ericzon, Shu-ichi Ikeda, W. David Lewis, et al. 2013. “Guideline of Transthyretin-Related Hereditary Amyloidosis for Clinicians.” *Orphanet Journal of Rare Diseases* 8 (February):31. https://doi.org/10.1186/1750-1172-8-31.

Andrew, Sarah M., and Julie A. Titus. 2001. “Fragmentation of Immunoglobulin G.” In *Current Protocols in Cell Biology*. John Wiley & Sons, Inc. http://onlinelibrary.wiley.com/doi/10.1002/0471143030.cb1604s17/abstract.

Anfinsen, C. B. 1973. “Principles That Govern the Folding of Protein Chains.” *Science (New York, N.Y.)* 181 (4096):223–30.

Asherie, Neer. 2004. “Protein Crystallization and Phase Diagrams.” *Methods*, Macromolecular Crystallization, 34 (3):266–72. https://doi.org/10.1016/j.ymeth.2004.03.028.

Atkins, Peter, and Julio de Paula. 2005. *Elements of Physical Chemistry*. 4 edition. New York: OUP.

Avadhanulu, M. N., and P. G. Kshirsagar. 2014. *A Textbook of Engineering Physics*. Tenth edition. S Chand.

Banypersad, Sanjay M., James C. Moon, Carol Whelan, Philip N. Hawkins, and Ashutosh D. Wechalekar. 2012. “Updates in Cardiac Amyloidosis:

A Review.” *Journal of the American Heart Association* 1 (2):e000364. https://doi.org/10.1161/JAHA.111.000364.

Baskakov, Ilia V., Giuseppe Legname, Stanley B. Prusiner, and Fred E. Cohen. 2001. “Folding of Prion Protein to Its Native α-Helical Conformation Is under Kinetic Control.” *Journal of Biological Chemistry* 276 (23):19687–90. https://doi.org/10.1074/jbc.C100180200.

Baures, P. W., V. B. Oza, S. A. Peterson, and J. W. Kelly. 1999. “Synthesis and Evaluation of Inhibitors of Transthyretin Amyloid Formation Based on the Non-Steroidal Anti-Inflammatory Drug, Flufenamic Acid.” *Bioorganic & Medicinal Chemistry* 7 (7):1339–47.

Benditt, E. P., N. Eriksen, and C. Berglund. 1970. “Congo Red Dichroism with Dispersed Amyloid Fibrils, an Extrinsic Cotton Effect.” *Proceedings of the National Academy of Sciences* 66 (4):1044–51.

Benhold, H. 1922. “Eigne Spezifische Amyloidfiirbung Mit Kongorot.” *Mi~ Rzch Med Wochenschr* 2:1537–1538.

Berg, Jeremy Mark, John L Tymoczko, and Lubert Stryer. 2012. *Biochemistry*. New York, N.Y: W.H. Freeman and Company.

Berg, Jeremy M., John L. Tymoczko, and Lubert Stryer. 2002. “Primary Structure: Amino Acids Are Linked by Peptide Bonds to Form Polypeptide Chains.” https://www.ncbi.nlm.nih.gov/books/NBK22364/.

Bickerstaff, M. C. M., M. Botto, W. L. Hutchinson, J. Herbert, G. A. Tennent, A. Bybee, D. A. Mitchell, et al. 1999. “Serum Amyloid P Component Controls Chromatin Degradation and Prevents Antinuclear Autoimmunity.” *Nature Medicine* 5 (6):694–97. https://doi.org/10.1038/9544.

Bodin, Karl, Stephan Ellmerich, Melvyn C. Kahan, Glenys A. Tennent, Andrzej Loesch, Janet A. Gilbertson, Winston L. Hutchinson, et al. 2010. “Antibodies to Human Serum Amyloid P Component Eliminate Visceral Amyloid Deposits.” *Nature* 468 (7320):93–97. https://doi.org/10.1038/nature09494.

Bolivar, F., R. L. Rodriguez, P. J. Greene, M. C. Betlach, H. L. Heyneker, H. W. Boyer, J. H. Crosa, and S. Falkow. 1977. “Construction and Characterization of New Cloning Vehicles. II. A Multipurpose Cloning System.” *Gene* 2 (2):95–113. https://doi.org/10.1016/0378-1119(77)90000-2.

Bolsover, Stephen R., Jeremy S. Hyams, Elizabeth A. Shephard, Hugh A. White, and Claudia G. Wiedemann. 2003. “Protein Structure.” In *Cell Biology*, 183–211. John Wiley & Sons, Inc. http://onlinelibrary.wiley.com/doi/10.1002/047146158X.ch9/summary.

Booth, David R., Margaret Sunde, Vittorio Bellotti, Carol V. Robinson, Winston L. Hutchinson, Paul E. Fraser, Philip N. Hawkins, et al. 1997. “Instability, Unfolding and Aggregation of Human Lysozyme Variants Underlying Amyloid Fibrillogenesis.” *Nature* 385 (6619):787–93. https://doi.org/10.1038/385787a0.

Botto, Marina, Philip N. Hawkins, Maria C. M. Bickerstaff, Jeff Herbert, Anne E. Bygrave, Alan Mcbride, Winston L. Hutchinson, Glenys A. Tennent, M. J. Walport, and Mark B. Pepys. 1997. “Amyloid Deposition Is Delayed in Mice with Targeted Deletion of the Serum Amyloid P Component Gene.” *ResearchGate* 3 (8):855–59. https://doi.org/10.1038/nm0897-855.

Bouchard, Mario, Jesús Zurdo, Ewan J. Nettleton, Christopher M. Dobson, and Carol V. Robinson. 2000. “Formation of Insulin Amyloid Fibrils Followed by FTIR Simultaneously with CD and Electron Microscopy.” *Protein Science* 9 (10):1960–67. https://doi.org/10.1110/ps.9.10.1960.

Bourgault, Steve, James P. Solomon, Natàlia Reixach, and Jeffery W. Kelly. 2011. “Sulfated Glycosaminoglycans Accelerate Transthyretin Amyloidogenesis by Quaternary Structural Conversion.” *Biochemistry* 50 (6):1001–15. https://doi.org/10.1021/bi101822y.

Bragg, W. H., and W. L. Bragg. 1913. “The Reflection of X-Rays by Crystals.” *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* 88 (605):428–38. https://doi.org/10.1098/rspa.1913.0040.

Branden, Carl, and John Tooze. 1991. *Introduction to Protein Structure*. Garland Pub.

Brange, Jens, Lennart Andersen, Erik D. Laursen, Giorgio Meyn, and Eigil Rasmussen. 1997. “Toward Understanding Insulin Fibrillation.” *Journal of Pharmaceutical Sciences* 86 (5):517–25. https://doi.org/10.1021/js960297s.

Bratanova-Tochkova, Troitza K., Haiying Cheng, Samira Daniel, Subhadra Gunawardana, Yi-Jia Liu, Jennifer Mulvaney-Musa, Thomas Schermerhorn, Susanne G. Straub, Hiroki Yajima, and Geoffrey W. G. Sharp. 2002. “Triggering and Augmentation Mechanisms, Granule Pools, and Biphasic Insulin Secretion.” *Diabetes* 51 Suppl 1 (February):S83-90.

Breathnach, S. M., H. Kofler, N. Sepp, J. Ashworth, D. Woodrow, M. B. Pepys, and H. Hintner. 1989. “Serum Amyloid P Component Binds to Cell Nuclei in Vitro and to in Vivo Deposits of Extracellular Chromatin in Systemic Lupus Erythematosus.” *The Journal of Experimental Medicine* 170 (4):1433–38. https://doi.org/10.1084/jem.170.4.1433.

Bright Focus (2015). *The Progression of Alzheimer’s Disease.* Retrieved from https://www.brightfocus.org/alzheimers/infographic/progression-alzheimers-disease

Bryngelson, Joseph D., José Nelson Onuchic, Nicholas D. Socci, and Peter G. Wolynes. 1995. “Funnels, Pathways, and the Energy Landscape of Protein Folding: A Synthesis.” *Proteins: Structure, Function, and Bioinformatics* 21 (3):167–95. https://doi.org/10.1002/prot.340210302.

Butterfield, D. Allan, and Debra Boyd-Kimball. 2004. “Amyloid Beta-peptide(1-42) Contributes to the Oxidative Stress and Neurodegeneration Found in Alzheimer Disease Brain.” *Brain Pathology (Zurich, Switzerland)* 14 (4):426–32.

Caccamo, Antonella, Salvatore Oddo, Michael C. Sugarman, Yama Akbari, and Frank M. LaFerla. 2005. “Age- and Region-Dependent Alterations in Abeta-Degrading Enzymes: Implications for Abeta-Induced Disorders.” *Neurobiology of Aging* 26 (5):645–54. https://doi.org/10.1016/j.neurobiolaging.2004.06.013.

Cai, Lisheng, Robert B. Innis, and Victor W. Pike. 2007. “Radioligand Development for PET Imaging of Beta-Amyloid (Abeta)--Current Status.” *Current Medicinal Chemistry* 14 (1):19–52.

Calkins, Evan. 1983. “Amyloid and Amyloidosis: Proceedings of the Third International Symposium on Amyloidosis, Povoa de Varzim, Portugal, September 23–28, 1979.” *Arthritis & Rheumatism* 26 (4):575–76. https://doi.org/10.1002/art.1780260429.

Catalano, Susan M., Elizabeth C. Dodson, Darrell A. Henze, Joseph G. Joyce, Grant A. Krafft, and Gene G. Kinney. 2006. “The Role of Amyloid-Beta Derived Diffusible Ligands (ADDLs) in Alzheimer’s Disease.” *Current Topics in Medicinal Chemistry* 6 (6):597–608.

Chayen, Naomi E. 1998. “Comparative Studies of Protein Crystallization by Vapour-Diffusion and Microbatch Techniques.” *Acta Crystallographica Section D Biological Crystallography* 54 (1):8–15. https://doi.org/10.1107/S0907444997005374.

“Chemistry of Protein Assays.” n.d. Accessed November 5, 2015. https://www.thermofisher.com/uk/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/chemistry-protein-assays.html.

Cheng, Yuan, Ted C. Judd, Michael D. Bartberger, James Brown, Kui Chen, Robert T. Fremeau, Dean Hickman, et al. 2011. “From Fragment Screening to in Vivo Efficacy: Optimization of a Series of 2-Aminoquinolines as Potent Inhibitors of Beta-Site Amyloid Precursor Protein Cleaving Enzyme 1 (BACE1).” *Journal of Medicinal Chemistry* 54 (16):5836–57. https://doi.org/10.1021/jm200544q.

Chu, Guang, Xuesi Wang, Tianrui Chen, Wen Xu, Yu Wang, Hongwei Song, and Yan Xu. 2015. “Chiral Electronic Transitions of YVO4:Eu3+ Nanoparticles in Cellulose Based Photonic Materials with Circularly Polarized Excitation.” *Journal of Materials Chemistry C* 3 (14):3384–90. https://doi.org/10.1039/C4TC02913E.

Cirrito, John R., Rashid Deane, Anne M. Fagan, Michael L. Spinner, Maia Parsadanian, Mary Beth Finn, Hong Jiang, et al. 2005. “P-Glycoprotein Deficiency at the Blood-Brain Barrier Increases Amyloid-Beta Deposition in an Alzheimer Disease Mouse Model.” *The Journal of Clinical Investigation* 115 (11):3285–90. https://doi.org/10.1172/JCI25247.

Ciszak, E., V. Cody, and J. R. Luft. 1992. “Crystal Structure Determination at 2.3-A Resolution of Human Transthyretin-3’,5’-dibromo-2’,4,4’,6-Tetrahydroxyaurone Complex.” *Proc.Natl.Acad.Sci.USA* 89 (April):6644–48. https://doi.org/10.1073/pnas.89.14.6644.

CK-12 (2017). *Humoral Immune Response.* Retrieved from https://www.ck12.org/book/CK-12-Biology-Concepts/section/13.51/

Coelho, T. 1996. “Familial Amyloid Polyneuropathy: New Developments in Genetics and Treatment.” *Current Opinion in Neurology* 9 (5):355–59.

Cohen, Samuel I. A., Michele Vendruscolo, Christopher M. Dobson, and Tuomas P. J. Knowles. 2012. “From Macroscopic Measurements to Microscopic Mechanisms of Protein Aggregation.” *Journal of Molecular Biology*, Amyloid Structure, Function, and Molecular Mechanisms (Part I), 421 (2–3):160–71. https://doi.org/10.1016/j.jmb.2012.02.031.

Colon, Wilfredo, and Jeffery W. Kelly. 1992. “Partial Denaturation of Transthyretin Is Sufficient for Amyloid Fibril Formation in Vitro.” *Biochemistry* 31 (36):8654–60. https://doi.org/10.1021/bi00151a036.

Cooper, Geoffrey M., and Robert E. Hausman. 2009. *The Cell: A Molecular Approach*. ASM Press.

D Thompson, M. B. Pepys. 2002. “The Structures of Crystalline Complexes of Human Serum Amyloid P Component with Its Carbohydrate Ligand, The Cyclic Pyruvate Acetal of Galactose.” *Journal of Molecular Biology* 320 (5):1081–86. https://doi.org/10.1016/S0022-2836(02)00514-4.

DeFronzo, R. A., E. Ferrannini, P. Zimmet, and George Alberti, eds. 2015. *International Textbook of Diabetes Mellitus*. 4th Revised edition edition. Chichester, West Sussex ; Hoboken, NJ: Wiley-Blackwell.

Delves, Peter, Seamus Martin, Dennis Burton, and Ivan Roitt. 2006. *Roitt’s Essential Immunology*. 11th Edition edition. Malden, Mass.: Wiley-Blackwell.

Dessau, Moshe A., and Yorgo Modis. 2011. “Protein Crystallization for X-Ray Crystallography.” *Journal of Visualized Experiments: JoVE*, no. 47. https://doi.org/10.3791/2285.

Divry, Paul. 1927. “Etude Histochimique Des Plaques Seniles.” *J Belge Neurol Psychiat* 27:643–657.

Döbeli, H., N. Draeger, G. Huber, P. Jakob, D. Schmidt, B. Seilheimer, D. Stüber, B. Wipf, and M. Zulauf. 1995. “A Biotechnological Method Provides Access to Aggregation Competent Monomeric Alzheimer’s 1-42 Residue Amyloid Peptide.” *Bio/Technology (Nature Publishing Company)* 13 (9):988–93.

Dobson, Christopher M. 2002. “Protein-Misfolding Diseases: Getting out of Shape.” *Nature* 418 (6899):729–30. https://doi.org/10.1038/418729a.

Dong, A., W. S. Caughey, and T. W. Du Clos. 1994. “Effects of Calcium, Magnesium, and Phosphorylcholine on Secondary Structures of Human C-Reactive Protein and Serum Amyloid P Component Observed by Infrared Spectroscopy.” *The Journal of Biological Chemistry* 269 (9):6424–30.

Drenth, Jan. 2013. *Principles of Protein X-Ray Crystallography*. Springer Science & Business Media.

D'Souza, Anita, Jason D. Theis, Julie A. Vrana, and Ahmet Dogan. 2014. "Pharmaceutical Amyloidosis Associated with Subcutaneous Insulin and Enfuvirtide Administration." *Amyloid: The International Journal of Experimental and Clinical Investigation: The Official Journal of the International Society of Amyloidosis* 21 (2):71–75. https://doi.org/10.3109/13506129.2013.876984.

Dubendorff, J. W., and F. W. Studier. 1991. "Controlling Basal Expression in an Inducible T7 Expression System by Blocking the Target T7 Promoter with Lac Repressor." *Journal of Molecular Biology* 219 (1):45–59.

Dubois, Bruno, Howard H. Feldman, Claudia Jacova, Steven T. Dekosky, Pascale Barberger-Gateau, Jeffrey Cummings, André Delacourte, et al. 2007. "Research Criteria for the Diagnosis of Alzheimer's Disease: Revising the NINCDS-ADRDA Criteria." *The Lancet. Neurology* 6 (8):734–46. https://doi.org/10.1016/S1474-4422(07)70178-3.

Ducruix, A., and R. Giegé, eds. 1992. *Crystallization of Nucleic Acids and Proteins: A Practical Approach*. IRL Press | Practical Approach Series 86.

Dumoulin, Mireille, Denis Canet, Alexander M. Last, Els Pardon, David B. Archer, Serge Muyldermans, Lode Wyns, et al. 2005. "Reduced Global Cooperativity Is a Common Feature Underlying the Amyloidogenicity of Pathogenic Lysozyme Mutations." *Journal of Molecular Biology* 346 (3):773–88. https://doi.org/10.1016/j.jmb.2004.11.020.

E. Bentley, William, Noushin Mirjalili, Dana Andersen, Robert H. Davis, and Dhinakar Kompala. 1990. "Plasmid-Encoded Protein: The Principal Factor in the 'Metabolic Burden' Associated with Recombinant Bacteria." *Biotechnology and Bioengineering* 35 (March):668–81. https://doi.org/10.1002/bit.260350704.

Endo, Justin O., Christoph Röcken, Sara Lamb, Ronald M. Harris, and Anneli R. Bowen. 2010. "Nodular Amyloidosis in a Diabetic Patient with Frequent Hypoglycemia: Sequelae of Repeatedly Injecting Insulin without Site Rotation." *Journal of the American Academy of Dermatology* 63 (6):e113-114. https://doi.org/10.1016/j.jaad.2010.03.001.

Falk, Rodney H. 2005. "Diagnosis and Management of the Cardiac Amyloidoses." *Circulation* 112 (13):2047–60. https://doi.org/10.1161/CIRCULATIONAHA.104.489187.

Fraser, P. E., J. T. Nguyen, W. K. Surewicz, and D. A. Kirschner. 1991. "pH-Dependent Structural Transitions of Alzheimer Amyloid Peptides." *Biophysical Journal* 60 (5):1190–1201. https://doi.org/10.1016/S0006-3495(91)82154-3.

G G Glenner, E. D. Eanes. 1975. "Beta-Pleated Sheet Fibrils - a Comparison of Native Amyloid with Synthetic Protein Fibrils. J Histochem Cytochem." *The Journal of Histochemistry and Cytochemistry : Official Journal of the Histochemistry Society* 22 (12):1141–58. https://doi.org/10.1177/22.12.1141.

Gailer, C., and M. Feigel. 1997. “Is the Parallel or Antiparallel Beta-Sheet More Stable? A Semiempirical Study.” *Journal of Computer-Aided Molecular Design* 11 (3):273–77.

Garman, E. F., and T. R. Schneider. 1997. “Macromolecular Cryocrystallography.” *Journal of Applied Crystallography* 30 (3):211–37. https://doi.org/10.1107/S0021889897002677.

Garman, Elspeth. 2003. “‘Cool’ Crystals: Macromolecular Cryocrystallography and Radiation Damage.” *Current Opinion in Structural Biology* 13 (5):545–51.

G Wagner, S G Hyberts, and T. F. Havel. 1992. “NMR Structure Determination in Solution: A Critique and Comparison with X-Ray Crystallography.” *Annual Review of Biophysics and Biomolecular Structure* 21 (1):167–98. https://doi.org/10.1146/annurev.bb.21.060192.001123.

Gertz, Morie A., and S. Vincent Rajkumar. 2013. “Primary Systemic Amyloidosis.” *Current Treatment Options in Oncology* 3 (3):261–71. https://doi.org/10.1007/s11864-002-0016-1.

Gething, Mary-Jane, and Joseph Sambrook. 1992. “Protein Folding in the Cell.” *Nature* 355 (6355):33–45. https://doi.org/10.1038/355033a0.

Gewurz, Henry, Xiao-Hui Zhang, and Thomas Franklin Lint. 1995. “Structure and Function of the Pentraxins.” *Current Opinion in Immunology* 7 (1):54–64. https://doi.org/10.1016/0952-7915(95)80029-8.

Goedert, Michel. 1993. “Tau Protein and the Neurofibrillary Pathology of Alzheimer’s Disease.” *Trends in Neurosciences* 16 (11):460–65. https://doi.org/10.1016/0166-2236(93)90078-Z.

Graham, M. M. 2011. “Clinical Molecular Imaging with Radiotracers: Current Status.” *Medical Principles and Practice* 21 (3):197–208. https://doi.org/10.1159/000333552.

Groenning, Minna. 2009. “Binding Mode of Thioflavin T and Other Molecular Probes in the Context of Amyloid Fibrils—current Status.” *Journal of Chemical Biology* 3 (1):1–18. https://doi.org/10.1007/s12154-009-0027-5.

Groenning, Minna, Mathias Norrman, James M. Flink, Marco van de Weert, Jens T. Bukrinsky, Gerd Schluckebier, and Sven Frokjaer. 2007. “Binding Mode of Thioflavin T in Insulin Amyloid Fibrils.” *Journal of Structural Biology* 159 (3):483–97. https://doi.org/10.1016/j.jsb.2007.06.004.

Gupta, Yashdeep, Gaurav Singla, and Rajiv Singla. 2015. “Insulin-Derived Amyloidosis.” *Indian Journal of Endocrinology and Metabolism* 19 (1):174–77. https://doi.org/10.4103/2230-8210.146879.

Hadley, Gina. 2007. “Basic Histology.” *Journal of Anatomy* 211 (3):412–13. https://doi.org/10.1111/j.1469-7580.2007.771_1.x.

Hamazaki, H. 1987. “Ca2+-Mediated Association of Human Serum Amyloid P Component with Heparan Sulfate and Dermatan Sulfate.” *Journal of Biological Chemistry* 262 (4):1456–60.

Hammarström, Per, R. Luke Wiseman, Evan T. Powers, and Jeffery W. Kelly. 2003. “Prevention of Transthyretin Amyloid Disease by Changing Protein Misfolding Energetics.” *Science (New York, N.Y.)* 299 (5607):713–16. https://doi.org/10.1126/science.1079589.

Hardy, John, and Dennis J. Selkoe. 2002. “The Amyloid Hypothesis of Alzheimer’s Disease: Progress and Problems on the Road to Therapeutics.” *Science (New York, N.Y.)* 297 (5580):353–56. https://doi.org/10.1126/science.1072994.

Harris, L. J., S. B. Larson, K. W. Hasel, and A. McPherson. 1996. “Refined Structure of an Intact IgG2a Monoclonal Antibody.” *Biochemistry* 36 (October):1581–97. https://doi.org/10.1021/bi962514+.

Hartl, F. Ulrich, and Manajit Hayer-Hartl. 2009. “Converging Concepts of Protein Folding in Vitro and in Vivo.” *Nature Structural & Molecular Biology* 16 (6):574–81. https://doi.org/10.1038/nsmb.1591.

Hartley, D. L., and J. F. Kane. 1988. “Properties of Inclusion Bodies from Recombinant Escherichia Coli.” *Biochemical Society Transactions* 16 (2):101–2.

Hawkins, P. N., J. P. Lavender, and M. B. Pepys. 1990. “Evaluation of Systemic Amyloidosis by Scintigraphy with 12I-Labeled Serum Amyloid P Component.” *The New England Journal of Medicine* 323 (8):508–13.

Hawkins, P. N., M. J. Myers, A. A. Epenetos, D. Caspi, and M. B. Pepys. 1988. “Specific Localization and Imaging of Amyloid Deposits in Vivo Using 123I-Labeled Serum Amyloid P Component.” *The Journal of Experimental Medicine* 167 (3):903–13. https://doi.org/10.1084/jem.167.3.903.

Hawkins, P. N., R. Wootton, and M. B. Pepys. 1991. “Metabolic Studies of Radioiodinated Serum Amyloid P Component in Normal Subjects and Patients with Systemic Amyloidosis.” In *Amyloid and Amyloidosis 1990*, 254–57. Springer, Dordrecht. https://link.springer.com/chapter/10.1007/978-94-011-3284-8_65.

Heilbrunn, L. V., Francis T. Ashton, Carl Feldherr, and Walter L. Wilson. 1958. “The Action of Insulin on Cells and Protoplasm.” *Biological Bulletin* 115 (3):459–70. https://doi.org/10.2307/1539109.

Heldt, Caryn L., Shuqi Zhang, and Georges Belfort. 2011. “Asymmetric Amyloid Fibril Elongation: A New Perspective on a Symmetric World.” *Proteins* 79 (1):92–98. https://doi.org/10.1002/prot.22861.

Hendrickson, W. A. 2000. “Synchrotron Crystallography.” *Trends in Biochemical Sciences* 25 (12):637–43.

Hicks, P. S., L. Saunero-Nava, T. W. Du Clos, and C. Mold. 1992. “Serum Amyloid P Component Binds to Histones and Activates the Classical Complement Pathway.” *The Journal of Immunology* 149 (11):3689–94.

Hohenester, Erhard, Winston L Hutchinson, Mark B Pepys, and Steve P Wood. 1997. “Crystal Structure of a Decameric Complex of Human Serum Amyloid P Component with Bound dAMP1.” *Journal of Molecular Biology* 269 (4):570–78. https://doi.org/10.1006/jmbi.1997.1075.

Holmgren, G, L Steen, O Suhr, B. -G Ericzon, C. -G Groth, O Andersen, B. G Wallin, et al. 1993. “Clinical Improvement and Amyloid Regression after Liver Transplantation in Hereditary Transthyretin Amyloidosis.” *The Lancet*, Originally published as Volume 1, Issue 8853, 341 (8853):1113–16. https://doi.org/10.1016/0140-6736(93)93127-M.

Hortschansky, Peter, Volker Schroeckh, Tony Christopeit, Giorgia Zandomeneghi, and Marcus Fändrich. 2005. “The Aggregation Kinetics of Alzheimer’s β-Amyloid Peptide Is Controlled by Stochastic Nucleation.” *Protein Science : A Publication of the Protein Society* 14 (7):1753–59. https://doi.org/10.1110/ps.041266605.

Horwich, Arthur L., Wayne A. Fenton, Eli Chapman, and George W. Farr. 2007. “Two Families of Chaperonin: Physiology and Mechanism.” *Annual Review of Cell and Developmental Biology* 23:115–45. https://doi.org/10.1146/annurev.cellbio.23.090506.123555.

Hutchinson, W. L., E. Hohenester, and M. B. Pepys. 2000. “Human Serum Amyloid P Component Is a Single Uncomplexed Pentamer in Whole Serum.” *MOL MED* 6 (6):482–93.

Immunoglobulin G Deficiency (2017). Retrieved from http://misc.medscape.com/pi/iphone/medscapeapp/html/A136897-business.html

István Likó, Marianna Mák. 2007. “Evidence for an Extended Interacting Surface between β-Amyloid and Serum Amyloid P Component.” *Neuroscience Letters* 412 (1):51–55. https://doi.org/10.1016/j.neulet.2006.10.052.

Ittner, Lars M., Yazi D. Ke, Fabien Delerue, Mian Bi, Amadeus Gladbach, Janet van Eersel, Heidrun Wölfing, et al. 2010. “Dendritic Function of Tau Mediates Amyloid-Beta Toxicity in Alzheimer’s Disease Mouse Models.” *Cell* 142 (3):387–97. https://doi.org/10.1016/j.cell.2010.06.036.

Jacobson, D. R., R. D. Pastore, R. Yaghoubian, I. Kane, G. Gallo, F. S. Buck, and J. N. Buxbaum. 1997. “Variant-Sequence Transthyretin (Isoleucine 122) in Late-Onset Cardiac Amyloidosis in Black Americans.” *The New England Journal of Medicine* 336 (7):466–73. https://doi.org/10.1056/NEJM199702133360703.

Janciauskiene, Sabina, Pablo García de Frutos, Erik Carlemalm, Björn Dahlbäck, and Sten Eriksson. 1995. “Inhibition of Alzheimer β-Peptide Fibril Formation by Serum Amyloid P Component.” *Journal of Biological Chemistry* 270 (44):26041–44. https://doi.org/10.1074/jbc.270.44.26041.

Janeway, Charles, Paul Travers, Mark Walport, and Mark Shlomchik. 2001. *Immunobiology: The Immune System in Health and Disease*. 5 edition. New York: Garland Science.

Jefferis, Roy. 2006. “A Sugar Switch for Anti-Inflammatory Antibodies.” *Nature Biotechnology* 24 (10):1230–31. https://doi.org/10.1038/nbt1006-1230.

Jensen, Ellen C. 2012. “The Basics of Western Blotting.” *The Anatomical Record: Advances in Integrative Anatomy and Evolutionary Biology* 295 (3):369–71. https://doi.org/10.1002/ar.22424.

Jenvey, M.C. (2006). *Structure Led Drug Design for the Pentraxins.* Unpublished clinical psychology doctoral thesis, Division of Biochemistry & Molecular Biology, Faculty of Medicine, Health & Life Sciences.

Jiang, Jianwen, Kondury Prasad, Eileen M. Lafer, and Rui Sousa. 2005. “Structural Basis of Interdomain Communication in the Hsc70 Chaperone.” *Molecular Cell* 20 (4):513–24. https://doi.org/10.1016/j.molcel.2005.09.028.

Jiménez, José L., Ewan J. Nettleton, Mario Bouchard, Carol V. Robinson, Christopher M. Dobson, and Helen R. Saibil. 2002. “The Protofilament Structure of Insulin Amyloid Fibrils.” *Proceedings of the National Academy of Sciences* 99 (14):9196–9201. https://doi.org/10.1073/pnas.142459399.

Jonas Emsley, Helen E. White. 1994. “Emsley, J. et Al. Structure of Pentameric Human Serum Amyloid P Component. Nature 367, 338-345.” *Nature* 367 (6461):338–45. https://doi.org/10.1038/367338a0.

Joseph T. Jarrett, Peter T. Lansbury. 1993. “Jarrett, J.T. & Lansbury, P.T., Jr. Seeding "one-Dimensional Crystallization" of Amyloid: A Pathogenic Mechanism in Alzheimer's Disease and Scrapie? Cell 73, 1055-1058.” *Cell* 73 (6):1055–58. https://doi.org/10.1016/0092-8674(93)90635-4.

Kagan, Herbert M., Nancy A. Hewitt, and Carl Franzblau. 1973. “A Microenvironmental Probe of Elastin. Properties of a Solubilized Congo Red-Elastin Complex.” *Biochimica et Biophysica Acta (BBA) - Protein Structure* 322 (2):258–68. https://doi.org/10.1016/0005-2795(73)90302-4.

Kamphuis, I. G., K. H. Kalk, M. B. Swarte, and J. Drenth. 1986. “Structure of Papain Refined at 1.65 A Resolution.” *J.Mol.Biol.* 179 (March):233–56. https://doi.org/10.1016/0022-2836(84)90467-4.

Kanekiyo, Takahisa, John R. Cirrito, Chia-Chen Liu, Mitsuru Shinohara, Jie Li, Dorothy R. Schuler, Motoko Shinohara, David M. Holtzman, and Guojun Bu. 2013. “Neuronal Clearance of Amyloid-β by Endocytic Receptor LRP1.” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 33 (49):19276–83. https://doi.org/10.1523/JNEUROSCI.3487-13.2013.

Kelényi, G. 1967. “On the Histochemistry of Azo Group-Free Thiazole Dyes.” *The Journal of Histochemistry and Cytochemistry: Official Journal of the Histochemistry Society* 15 (3):172–80.

Kelly, Sharon M., Thomas J. Jess, and Nicholas C. Price. 2005. “How to Study Proteins by Circular Dichroism.” *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* 1751 (2):119–39. https://doi.org/10.1016/j.bbapap.2005.06.005.

Kendrew, J. C., G. Bodo, H. M. Dintzis, R. G. Parrish, H. Wyckoff, and D. C. Phillips. 1958. “A Three-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis.” *Nature* 181 (4610):662–66. https://doi.org/10.1038/181662a0.

Khatami, Mahin, ed. 2012. *Inflammatory Diseases - Immunopathology, Clinical and Pharmacological Bases*. InTech. http://www.intechopen.com/books/inflammatory-diseases-immunopathology-clinical-and-pharmacological-bases.

Khurana, R., V. N. Uversky, L. Nielsen, and A. L. Fink. 2001. “Is Congo Red an Amyloid-Specific Dye?” *The Journal of Biological Chemistry* 276 (25):22715–21. https://doi.org/10.1074/jbc.M011499200.

Kihara, Takeshi, and Shun Shimohama. 2004. “Alzheimer’s Disease and Acetylcholine Receptors.” *Acta Neurobiologiae Experimentalis* 64 (1):99–105.

Kimple, Michelle E., Allison L. Brill, and Renee L. Pasker. 2013. “Overview of Affinity Tags for Protein Purification.” *Current Protocols in Protein Science / Editorial Board, John E. Coligan ... [et Al.]* 73 (September):Unit-9.9. https://doi.org/10.1002/0471140864.ps0909s73.

Kinoshita, C. M., A. T. Gewurz, J. N. Siegel, S. C. Ying, T. E. Hugli, J. E. Coe, R. K. Gupta, R. Huckman, and H. Gewurz. 1992. “A Protease-Sensitive Site in the Proposed Ca(2+)-Binding Region of Human Serum Amyloid P Component and Other Pentraxins.” *Protein Science: A Publication of the Protein Society* 1 (6):700–709. https://doi.org/10.1002/pro.5560010602.

Klunk, W. E., J. W. Pettegrew, and D. J. Abraham. 1989. “Quantitative Evaluation of Congo Red Binding to Amyloid-like Proteins with a Beta-Pleated Sheet Conformation.” *The Journal of Histochemistry and Cytochemistry: Official Journal of the Histochemistry Society* 37 (8):1273–81.

Kolstoe, S. E., B. H. Ridha, V. Bellotti, N. Wang, C. V. Robinson, S. J. Crutch, G. Keir, et al. 2009. “Molecular Dissection of Alzheimer’s Disease Neuropathology by Depletion of Serum Amyloid P Component.” *Proceedings of the National Academy of Sciences* 106 (18):7619–23. https://doi.org/10.1073/pnas.0902640106.

Kolstoe, S.E. (2005) *Ligand Binding to Pentraxins.* Unpublished clinical psychology doctoral thesis, Department of X-ray Crystallography, University of Southampton.

Korpimäki, Teemu, Jussi Kurittu, and Matti Karp. 2003. “Surprisingly Fast Disappearance of Beta-Lactam Selection Pressure in Cultivation as Detected with Novel Biosensing Approaches.” *Journal of Microbiological Methods* 53 (1):37–42.

Krebs, M. R. H., E. H. C. Bromley, and A. M. Donald. 2005. “The Binding of Thioflavin-T to Amyloid Fibrils: Localisation and Implications.” *Journal of Structural Biology* 149 (1):30–37. https://doi.org/10.1016/j.jsb.2004.08.002.

Kugimiya, Tomoe, Hirofumi Jono, Shiori Saito, Toru Maruyama, Daisuke Kadowaki, Yohei Misumi, Yoshinobu Hoshii, et al. 2011. “Loss of Functional Albumin Triggers Acceleration of Transthyretin Amyloid Fibril Formation in Familial Amyloidotic Polyneuropathy.” *Laboratory Investigation* 91 (8):1219–28. https://doi.org/10.1038/labinvest.2011.71.

Ladd, Mark, and Rex Palmer. 2014. *Structure Determination by X-Ray Crystallography: Analysis by X-Rays and Neutrons*. Springer Science & Business Media.

Lashuel, Hilal A., Christine Wurth, Linda Woo, and Jeffery W. Kelly. 1999. “The Most Pathogenic Transthyretin Variant, L55P, Forms Amyloid Fibrils under Acidic Conditions and Protofilaments under Physiological Conditions.” *Biochemistry* 38 (41):13560–73. https://doi.org/10.1021/bi991021c.

Lee, Eun Kyung, Jin Ha Hwang, Dong Yeon Shin, Dae Ihn Kim, and Yung Joon Yoo. 2005. “Production of Recombinant Amyloid-Beta Peptide 42 as an Ubiquitin Extension.” *Protein Expression and Purification* 40 (1):183–89. https://doi.org/10.1016/j.pep.2004.12.014.

Lesk, Arthur. 2010a. *Introduction to Protein Science: Architecture, Function, and Genomics*. 2 edition. Oxford ; New York: OUP Oxford.

———. 2010b. *Introduction to Protein Science: Architecture, Function, and Genomics*. OUP Oxford.

Lesné, Sylvain, Ming Teng Koh, Linda Kotilinek, Rakez Kayed, Charles G. Glabe, Austin Yang, Michela Gallagher, and Karen H. Ashe. 2006. “A Specific Amyloid-Beta Protein Assembly in the Brain Impairs Memory.” *Nature* 440 (7082):352–57. https://doi.org/10.1038/nature04533.

LeVine, H. 1997. “Stopped-Flow Kinetics Reveal Multiple Phases of Thioflavin T Binding to Alzheimer Beta (1-40) Amyloid Fibrils.” *Archives of Biochemistry and Biophysics* 342 (2):306–16. https://doi.org/10.1006/abbi.1997.0137.

LeVine, Harry. 1995. “Thioflavine T Interaction with Amyloid β-Sheet Structures.” *Amyloid* 2 (1):1–6.

Lilius, G., M. Persson, L. Bülow, and K. Mosbach. 1991. “Metal Affinity Precipitation of Proteins Carrying Genetically Attached Polyhistidine Affinity Tails.” *European Journal of Biochemistry* 198 (2):499–504. https://doi.org/10.3109/13506129509031881.

Liljas, A. (2009). *Textbook of Structural Biology* (2nd ed.). World Scientific Publishing.

Lin, Zong, Damian Madan, and Hays S. Rye. 2008. “GroEL Stimulates Protein Folding through Forced Unfolding.” *ResearchGate* 15 (3):303–11. https://doi.org/10.1038/nsmb.1394.

Lindley, P. F. 1999. “Macromolecular Crystallography with a Third-Generation Synchrotron Source.” *Acta Crystallographica. Section D, Biological Crystallography* 55 (Pt 10):1654–62.

Lodish, Harvey, Arnold Berk, and Chris A. Kaiser. 2007. *Molecular Cell Biology*. 6 edition. New York: W.H.Freeman.

Lomakin, A., D. S. Chung, G. B. Benedek, D. A. Kirschner, and D. B. Teplow. 1996. “On the Nucleation and Growth of Amyloid Beta-Protein Fibrils: Detection of Nuclei and Quantitation of Rate Constants.” *Proceedings of the National Academy of Sciences* 93 (3):1125–29.

Lonsdale-Eccles, A. A., P. Gonda, J. A. Gilbertson, and A. E. Haworth. 2009. “Localized Cutaneous Amyloid at an Insulin Injection Site.” *Clinical and Experimental Dermatology* 34 (8):e1027-1028. https://doi.org/10.1111/j.1365-2230.2009.03711.x.

Lowe, G. 1970. “The Structure and Mechanism of Action of Papain.” *Philosophical Transactions of the Royal Society of London B: Biological Sciences* 257 (813):237–48. https://doi.org/10.1098/rstb.1970.0023.

Luo, Yi, Brad Bolon, Michael A. Damore, Dan Fitzpatrick, Hantao Liu, Jianhua Zhang, Qiao Yan, Robert Vassar, and Martin Citron. 2003. “BACE1 (Beta-Secretase) Knockout Mice Do Not Acquire Compensatory Gene Expression Changes or Develop Neural Lesions over Time.” *Neurobiology of Disease* 14 (1):81–88.

Ma, Buyong, and Ruth Nussinov. 2012. “Selective Molecular Recognition in Amyloid Growth and Transmission and Cross-Species Barriers.” *Journal of Molecular Biology*, Amyloid Structure, Function, and Molecular Mechanisms (Part I), 421 (2–3):172–84. https://doi.org/10.1016/j.jmb.2011.11.023.

Madej, Thomas, Christopher J. Lanczycki, Dachuan Zhang, Paul A. Thiessen, Renata C. Geer, Aron Marchler-Bauer, and Stephen H. Bryant. 2014. “MMDB and VAST+: Tracking Structural Similarities between Macromolecular Complexes.” *Nucleic Acids Research* 42 (D1):D297–303. https://doi.org/10.1093/nar/gkt1208.

Mayer, Matthias P. 2010. “Gymnastics of Molecular Chaperones.” *Molecular Cell* 39 (3):321–31. https://doi.org/10.1016/j.molcel.2010.07.012.

McPherson, A. 1990. “Current Approaches to Macromolecular Crystallization.” *European Journal of Biochemistry / FEBS* 189 (1):1–23.

Minton, N. P. 1984. “Improved Plasmid Vectors for the Isolation of Translational Lac Gene Fusions.” *Gene* 31 (1–3):269–73.

Miura, Takashi, Chiaki Yamamiya, Miho Sasaki, Kiyoko Suzuki, and Hideo Takeuchi. 2002. “Binding Mode of Congo Red to Alzheimer’s Amyloid β-Peptide Studied by UV Raman Spectroscopy.” *Journal of Raman Spectroscopy* 33 (7):530–35. https://doi.org/10.1002/jrs.869.

Moreira, Paula I., Cristina Carvalho, Xiongwei Zhu, Mark A. Smith, and George Perry. 2010. “Mitochondrial Dysfunction Is a Trigger of Alzheimer’s Disease Pathophysiology.” *Biochimica Et Biophysica Acta* 1802 (1):2–10. https://doi.org/10.1016/j.bbadis.2009.10.006.

Morris, A. L., M. W. MacArthur, E. G. Hutchinson, and J. M. Thornton. 1992. “Stereochemical Quality of Protein Structure Coordinates.” *Proteins* 12 (4):345–64. https://doi.org/10.1002/prot.340120407.

Murrell, J. R., R. G. Schoner, J. J. Liepnieks, H. N. Rosen, A. C. Moses, and M. D. Benson. 1992. “Production and Functional Analysis of Normal and Variant Recombinant Human Transthyretin Proteins.” *The Journal of Biological Chemistry* 267 (23):16595–600.

Nagase, Terumasa, Keiichi Iwaya, Yoshiki Iwaki, Fumio Kotake, Ryuji Uchida, Tsunao Oh-i, Hidenori Sekine, et al. 2014. “Insulin-Derived Amyloidosis and Poor Glycemic Control: A Case Series.” *The American Journal of Medicine* 127 (5):450–54. https://doi.org/10.1016/j.amjmed.2013.10.029.

Nagata-Uchiyama, M., M. Yaguchi, Y. Hirano, and T. Ueda. 2007. “Expression and Purification of Uniformly (15)N-Labeled Amyloid Beta Peptide 1-40 in Escherichia Coli.” *Protein and Peptide Letters* 14 (8):788–92.

Nelson, Laerke M., Luit Penninga, Kaare Sander, Peter B. Hansen, Gerda E. Villadsen, Allan Rasmussen, and Finn Gustafsson. 2013. “Long-Term Outcome in Patients Treated with Combined Heart and Liver Transplantation for Familial Amyloidotic Cardiomyopathy.” *Clinical Transplantation* 27 (2):203–9. https://doi.org/10.1111/ctr.12053.

Nilsson, Melanie R. 2004. “Techniques to Study Amyloid Fibril Formation in Vitro.” *Methods (San Diego, Calif.)* 34 (1):151–60. https://doi.org/10.1016/j.ymeth.2004.03.012.

Nordberg, Agneta. 2008. “Amyloid Plaque Imaging in Vivo: Current Achievement and Future Prospects.” *European Journal of Nuclear Medicine and Molecular Imaging* 35 Suppl 1 (March):S46-50. https://doi.org/10.1007/s00259-007-0700-2.

Nunomura, Akihiko, Rudy J. Castellani, Hyoung-Gon Lee, Paula I. Moreira, Xiongwei Zhu, George Perry, and Mark A. Smith. 2006. “Neuropathology in Alzheimer’s Disease: Awaking from a Hundred-Year-Old Dream.” *Science of Aging Knowledge Environment: SAGE KE* 2006 (8):pe10. https://doi.org/10.1126/sageke.2006.8.pe10.

O’Hara, B. P., S. P. Wood, G. Oliva, H. E. White, and M. B. Pepys. 1988. “Crystallizations of Human Serum Amyloid P Component (SAP).” *Journal of Crystal Growth* 90 (1–3):209–12. https://doi.org/10.1016/0022-0248(88)90317-X.

O'Hara, Carl J., and Rodney H. Falk. 2003. "The Diagnosis and Typing of Cardiac Amyloidosis." *Amyloid* 10 (2):127–29. https://doi.org/10.3109/13506120309041735.

Olsson, Fredrik, Staffan Schmidt, Veit Althoff, Lisa M. Munter, Shaobo Jin, Susanne Rosqvist, Urban Lendahl, Gerd Multhaup, and Johan Lundkvist. 2014. "Characterization of Intermediate Steps in Amyloid Beta (Aβ) Production under Near-Native Conditions." *Journal of Biological Chemistry* 289 (3):1540–50. https://doi.org/10.1074/jbc.M113.498246.

Olsson, Malin, Jenni Jonasson, Kristina Cederquist, and Ole B. Suhr. 2014. "Frequency of the Transthyretin Val30Met Mutation in the Northern Swedish Population." *Amyloid* 21 (1):18–20. https://doi.org/10.3109/13506129.2013.860027.

O'Nuallain, Brian, and Ronald Wetzel. 2002. "Conformational Abs Recognizing a Generic Amyloid Fibril Epitope." *Proceedings of the National Academy of Sciences* 99 (3):1485–90. https://doi.org/10.1073/pnas.022662599.

Overview of Western Blotting (2017). *Pierce Protein Methods*. Retrieved from https://www.thermofisher.com/uk/en/home/life-science/protein-biology/protein-biology-learning-center/protein-biology-resource-library/pierce-protein-methods/overview-western-blotting.html

Ozawa, Daisaku, Ryo Nomura, P. Patrizia Mangione, Kazuhiro Hasegawa, Tadakazu Okoshi, Riccardo Porcari, Vittorio Bellotti, and Hironobu Naiki. 2016. "Multifaceted Anti-Amyloidogenic and pro-Amyloidogenic Effects of C-Reactive Protein and Serum Amyloid P Component in Vitro." *Scientific Reports* 6 (July):29077. https://doi.org/10.1038/srep29077.

Palaninathan, Satheesh K., Nilofar N. Mohamedmohaideen, William C. Snee, Jeffery W. Kelly, and James C. Sacchettini. 2008. "Structural Insight into pH-Induced Conformational Changes within the Native Human Transthyretin Tetramer." *Journal of Molecular Biology* 382 (5):1157–67. https://doi.org/10.1016/j.jmb.2008.07.029.

Paul, William E. 2008. *Fundamental Immunology*. Lippincott Williams & Wilkins.

Pepys, M. B. 2001. "Pathogenesis, Diagnosis and Treatment of Systemic Amyloidosis." *PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY B-BIOLOGICAL SCIENCES* 356 (1406):203–10.

Pepys, M. B., and P. J. Butler. 1987. "Serum Amyloid P Component Is the Major Calcium-Dependent Specific DNA Binding Protein of the Serum." *Biochemical and Biophysical Research Communications* 148 (1):308–13.

Pepys, M. B, A. C Dash, E. A Munn, A Feinstein, Martha Skinner, A. S Cohen, H Gewurz, A. P Osmand, and R. H Painter. 1977. "ISOLATION OF AMYLOID P COMPONENT (PROTEIN AP) FROM NORMAL SERUM AS A CALCIUM-DEPENDENT BINDING PROTEIN." *The*

Lancet, Originally published as Volume 1, Issue 8020, 309 (8020):1029–31. https://doi.org/10.1016/S0140-6736(77)91260-0.

Pepys, M. B., J. Herbert, W. L. Hutchinson, G. A. Tennent, H. J. Lachmann, J. R. Gallimore, L. B. Lovat, et al. 2002. "Targeted Pharmacological Depletion of Serum Amyloid P Component for Treatment of Human Amyloidosis." *Nature* 417 (6886):254–59. https://doi.org/10.1038/417254a.

Pepys, Mark B. 2006. "Amyloidosis." *Annual Review of Medicine* 57 (1):223–41. https://doi.org/10.1146/annurev.med.57.121304.131243.

Pettersson, Tom, and Yrjö T. Konttinen. 2010. "Amyloidosis—Recent Developments." *Seminars in Arthritis and Rheumatism* 39 (5):356–68. https://doi.org/10.1016/j.semarthrit.2008.09.001.

Planté-Bordeneuve, Violaine, and Gerard Said. 2011. "Familial Amyloid Polyneuropathy." *The Lancet Neurology* 10 (12):1086–97. https://doi.org/10.1016/S1474-4422(11)70246-0.

Protein Purification (2017). *Custom Affinity Chromatography Service.* Retrieved from https://www.creative-biostructure.com/custom-affinity-chromatography-service-257.htm

Purkey, Hans E., Satheesh K. Palaninathan, Kathleen C. Kent, Craig Smith, Stephen H. Safe, James C. Sacchettini, and Jeffery W. Kelly. 2004. "Hydroxylated Polychlorinated Biphenyls Selectively Bind Transthyretin in Blood and Inhibit Amyloidogenesis: Rationalizing Rodent PCB Toxicity." *Chemistry & Biology* 11 (12):1719–28. https://doi.org/10.1016/j.chembiol.2004.10.009.

Qiu, Chengxuan, Miia Kivipelto, and Eva von Strauss. 2009. "Epidemiology of Alzheimer's Disease: Occurrence, Determinants, and Strategies toward Intervention." *Dialogues in Clinical Neuroscience* 11 (2):111–28.

Qiu, Wei Qiao, and Marshal F. Folstein. 2006. "Insulin, Insulin-Degrading Enzyme and Amyloid-Beta Peptide in Alzheimer's Disease: Review and Hypothesis." *Neurobiology of Aging* 27 (2):190–98. https://doi.org/10.1016/j.neurobiolaging.2005.01.004.

Ramachandran, G. N., C. Ramakrishnan, and V. Sasisekharan. 1963. "Stereochemistry of Polypeptide Chain Configurations." *Journal of Molecular Biology* 7 (July):95–99.

Razavi, Hossein, Satheesh K. Palaninathan, Evan T. Powers, R. Luke Wiseman, Hans E. Purkey, Nilofar N. Mohamedmohaideen, Songpon Deechongkit, et al. 2003. "Benzoxazoles as Transthyretin Amyloid Fibril Inhibitors: Synthesis, Evaluation, and Mechanism of Action." *Angewandte Chemie (International Ed. in English)* 42 (24):2758–61. https://doi.org/10.1002/anie.200351179.

Real de Asúa, Diego, Ramón Costa, Jose María Galván, María Teresa Filigheddu, Davinia Trujillo, and Julen Cadiñanos. 2014. "Systemic AA Amyloidosis: Epidemiology, Diagnosis, and Management." *Clinical*

Epidemiology 6 (October):369–77. https://doi.org/10.2147/CLEP.S39981.

Rhodes, Gale. 2012. *Crystallography Made Crystal Clear: A Guide for Users of Macromolecular Models*. Academic Press.

Richards, Duncan B., Louise M. Cookson, Alienor C. Berges, Sharon V. Barton, Thirusha Lane, James M. Ritter, Marianna Fontana, et al. 2015. "Therapeutic Clearance of Amyloid by Antibodies to Serum Amyloid P Component." *New England Journal of Medicine* 373 (12):1106–14. https://doi.org/10.1056/NEJMoa1504942.

Ridge, Perry G., Mark T. W. Ebbert, and John S. K. Kauwe. 2013. "Genetics of Alzheimer's Disease." *BioMed Research International* 2013:254954. https://doi.org/10.1155/2013/254954.

Rochet, Jean-Christophe, Kelly A. Conway, and Peter T. Lansbury. 2000. "Inhibition of Fibrillization and Accumulation of Prefibrillar Oligomers in Mixtures of Human and Mouse α-Synuclein†." *Biochemistry* 39 (35):10619–26. https://doi.org/10.1021/bi001315u.

Rock, Kenneth L., and Hajime Kono. 2008. "The Inflammatory Response to Cell Death." *Annual Review of Pathology* 3:99–126. https://doi.org/10.1146/annurev.pathmechdis.3.121806.151456.

Rosano, Germán L., and Eduardo A. Ceccarelli. 2014. "Recombinant Protein Expression in Escherichia Coli: Advances and Challenges." *Frontiers in Microbiology* 5 (April). https://doi.org/10.3389/fmicb.2014.00172.

Rubio, Nadina, Paul M. Sharp, Miriam Rits, Kamyar Zahedi, and Alexander S. Whitehead. 1993. "Structure, Expression, and Evolution of Guinea Pig Serum Amyloid P Component and C-Reactive Protein." *Journal of Biochemistry* 113 (3):277–84.

Saeed, S. M., and G. Fine. 1967. "Thioflavin-T for Amyloid Detection." *American Journal of Clinical Pathology* 47 (5):588–93.

Sathish, H. A., P. Kaul, and V. Prakash. 2000. "Influence of Metal Ions on Structure and Catalytic Activity of Papain." *Indian Journal of Biochemistry & Biophysics* 37 (1):18–27.

Schenk, Dale, Robin Barbour, Whitney Dunn, Grace Gordon, Henry Grajeda, Teresa Guido, Kang Hu, et al. 1999. "Immunization with Amyloid-β Attenuates Alzheimer-Disease-like Pathology in the PDAPP Mouse." *Nature* 400 (6740):173–77. https://doi.org/10.1038/22124.

Schmidt, H. H., T. D. Warner, K. Ishii, H. Sheng, and F. Murad. 1992. "Insulin Secretion from Pancreatic B Cells Caused by L-Arginine-Derived Nitrogen Oxides." *Science* 255 (5045):721–23. https://doi.org/10.1126/science.1371193.

Schröder, D., and H. Zühlke. 1982. "Gene Technology, Characterization of Insulin Gene and the Relationship to Diabetes Research." *Endokrinologie* 79 (2):197–209.

Schroeder Jr., Harry W., and Lisa Cavacini. 2010. "Structure and Function of Immunoglobulins." *Journal of Allergy and Clinical Immunology*, 2010

Primer on Allergic and Immunologic Diseases, 125 (2, Supplement 2):S41–52. https://doi.org/10.1016/j.jaci.2009.09.046.

Selkoe, Dennis J. 2003. “Folding Proteins in Fatal Ways.” *Nature* 426 (6968):900–904. https://doi.org/10.1038/nature02264.

Serpell, Louise C, Margaret Sunde, Merrill D Benson, Glenys A Tennent, Mark B Pepys, and Paul E Fraser. 2000. “The Protofilament Substructure of Amyloid fibrils1.” *Journal of Molecular Biology* 300 (5):1033–39. https://doi.org/10.1006/jmbi.2000.3908.

Sezonov, Guennadi, Danièle Joseleau-Petit, and Richard D’Ari. 2007. “Escherichia Coli Physiology in Luria-Bertani Broth.” *Journal of Bacteriology* 189 (23):8746–49. https://doi.org/10.1128/JB.01368-07.

Shahnawaz, Mohammad, Arjun Thapa, and Il-Seon Park. 2007. “Stable Activity of a Deubiquitylating Enzyme (Usp2-Cc) in the Presence of High Concentrations of Urea and Its Application to Purify Aggregation-Prone Peptides.” *Biochemical and Biophysical Research Communications* 359 (3):801–5. https://doi.org/10.1016/j.bbrc.2007.05.186.

Shankar, Ganesh M, and Dominic M Walsh. 2009. “Alzheimer’s Disease: Synaptic Dysfunction and Aβ.” *Molecular Neurodegeneration* 4 (November):48. https://doi.org/10.1186/1750-1326-4-48.

Sharp, Kim A. 1991. “The Hydrophobic Effect.” *Current Opinion in Structural Biology* 1 (2):171–74. https://doi.org/10.1016/0959-440X(91)90057-Z.

Shikama, Yasuyoshi, Jun-Ichi Kitazawa, Nobuo Yagihashi, Osamu Uehara, Yuuji Murata, Nobuhisa Yajima, Ryuichi Wada, and Soroku Yagihashi. 2010. “Localized Amyloidosis at the Site of Repeated Insulin Injection in a Diabetic Patient.” *Internal Medicine (Tokyo, Japan)* 49 (5):397–401.

Singh, Devraj. 2010. *Fundamentals of Optics*. 1 edition. New Delhi: Prentice-Hall of India Pvt.Ltd.

Sipe, Jean D., ed. 2006. *Amyloid Proteins: The Beta Sheet Conformation and Disease*. 2 Volumes edition. Weinheim ; Cambridge: Wiley VCH.

Sipe, Jean D., Merrill D. Benson, Joel N. Buxbaum, Shu-ichi Ikeda, Giampaolo Merlini, Maria J. M. Saraiva, and Per Westermark. 2016. “Amyloid Fibril Proteins and Amyloidosis: Chemical Identification and Clinical Classification International Society of Amyloidosis 2016 Nomenclature Guidelines.” *Amyloid* 23 (4):209–13. https://doi.org/10.1080/13506129.2016.1257986.

Skarzynski, Tadeusz. 2013. “Collecting Data in the Home Laboratory: Evolution of X-Ray Sources, Detectors and Working Practices.” *Acta Crystallographica Section D: Biological Crystallography* 69 (Pt 7):1283–88. https://doi.org/10.1107/S0907444913013619.

Skinner, Martha, and Alan S. Cohen. 1988. “[46] Amyloid P Component.” In , edited by BT - Methods in Enzymology, 163:523–36. Immunochemical Techniques Part M: Chemotaxis and Inflammation. Academic Press. http://www.sciencedirect.com/science/article/pii/0076687988630485.

Smyth, M S, and J H J Martin. 2000. “X Ray Crystallography.” *Molecular Pathology* 53 (1):8–14.

Sokolowski, Fabian, Andreas Johannes Modler, Ralf Masuch, Dietrich Zirwer, Michael Baier, Gudrun Lutsch, David Alan Moss, Klaus Gast, and Dieter Naumann. 2003. “Formation of Critical Oligomers Is a Key Event during Conformational Transition of Recombinant Syrian Hamster Prion Protein.” *Journal of Biological Chemistry* 278 (42):40481–92. https://doi.org/10.1074/jbc.M304391200.

Sørensen, Hans Peter, and Kim Kusk Mortensen. 2005. “Advanced Genetic Strategies for Recombinant Protein Expression in Escherichia Coli.” *Journal of Biotechnology* 115 (2):113–28. https://doi.org/10.1016/j.jbiotec.2004.08.004.

Sørensen, I. J., O. Andersen, E. H. Nielsen, and S. E. Svehag. 1995. “Native Human Serum Amyloid P Component Is a Single Pentamer.” *Scandinavian Journal of Immunology* 41 (3):263–67.

Sørensen, I. J., E. Holm Nielsen, L. Schrøder, A. Voss, L. Horváth, and S. E. Svehag. 2000. “Complexes of Serum Amyloid P Component and DNA in Serum from Healthy Individuals and Systemic Lupus Erythematosus Patients.” *Journal of Clinical Immunology* 20 (6):408–15.

Soria, Bernat, Ivan Quesada, Ana B. Ropero, José A. Pertusa, Franz Martín, and Angel Nadal. 2004. “Novel Players in Pancreatic Islet Signaling: From Membrane Receptors to Nuclear Channels.” *Diabetes* 53 Suppl 1 (February):S86-91.

Srinivasan, N, Helen E White, Jonas Emsley, Steve P Wood, Mark B Pepys, and Tom L Blundell. 1994. “Comparative Analyses of Pentraxins: Implications for Protomer Assembly and Ligand Binding.” *Structure* 2 (11):1017–27. https://doi.org/10.1016/S0969-2126(94)00105-7.

Störkel, S., H. M. Schneider, H. Müntefering, and S. Kashiwagi. 1983. “Iatrogenic, Insulin-Dependent, Local Amyloidosis.” *Laboratory Investigation; a Journal of Technical Methods and Pathology* 48 (1):108–11.

Støy, Julie, Emma L. Edghill, Sarah E. Flanagan, Honggang Ye, Veronica P. Paz, Anna Pluzhnikov, Jennifer E. Below, et al. 2007. “Insulin Gene Mutations as a Cause of Permanent Neonatal Diabetes.” *Proceedings of the National Academy of Sciences* 104 (38):15040–44. https://doi.org/10.1073/pnas.0707291104.

Studier, F. W., and B. A. Moffatt. 1986. “Use of Bacteriophage T7 RNA Polymerase to Direct Selective High-Level Expression of Cloned Genes.” *Journal of Molecular Biology* 189 (1):113–30.

Suhr, Ole B., Isabel M. Conceição, Onur N. Karayal, Francine S. Mandel, Pedro E. Huertas, and Bo-Göran Ericzon. 2014. “Post Hoc Analysis of Nutritional Status in Patients with Transthyretin Familial Amyloid Polyneuropathy: Impact of Tafamidis.” *Neurology and Therapy* 3 (2):101–12. https://doi.org/10.1007/s40120-014-0023-8.

Systems Immunology (2017). *Laboratory for Systems and Synthetic Technology.* Retrieved from https://www.bsse.ethz.ch/lsi/research/systems-immunology.html

Tabaton, Massimo, Xiongwei Zhu, George Perry, Mark A. Smith, and Luca Giliberto. 2010. “Signaling Effect of Amyloid-β42 on the Processing of AβPP.” *Experimental Neurology* 221 (1):18–25. https://doi.org/10.1016/j.expneurol.2009.09.002.

Tahrin Mahmood, Ping-Chang Yang. 2012. “Western Blot: Techique, Theory, and Trouble Shooting.” *North American Journal of Medical Sciences* 4 (9):429–34. https://doi.org/10.4103/1947-2714.100998.

Takahiro Tajiri, Yukio Ando. 2002. “Amyloid Formation in Rat Transthyretin: Effect of Oxidative Stress.” *Clinica Chimica Acta; International Journal of Clinical Chemistry* 323 (1–2):129–37. https://doi.org/10.1016/S0009-8981(02)00179-1.

Tantau, Alina, Mihaela Laszlo, and Istvan Laszlo. 2015. “Transthyretin Amyloidosis: An over Review.” *Cardiovascular Regenerative Medicine* 2 (0). https://doi.org/10.14800/crm.952.

Tennent, G. A., L. B. Lovat, and M. B. Pepys. 1995. “Serum Amyloid P Component Prevents Proteolysis of the Amyloid Fibrils of Alzheimer Disease and Systemic Amyloidosis.” *Proceedings of the National Academy of Sciences* 92 (10):4299–4303.

Thapa, Arjun, Md Shahnawaz, Pratap Karki, Giri Raj Dahal, Md Golam Sharoar, Song Yub Shin, Jung Sup Lee, Byungyun Cho, and Il-Seon Park. 2008. “Purification of Inclusion Body-Forming Peptides and Proteins in Soluble Form by Fusion to Escherichia Coli Thermostable Proteins.” *BioTechniques* 44 (6):787–96. https://doi.org/10.2144/000112728.

“The Progression of Alzheimer’s Disease.” 2015. BrightFocus Foundation. July 2, 2015. http://www.brightfocus.org/alzheimers/infographic/progression-alzheimers-disease.

Vassar, P. S., and C. F. Culling. 1959. “Fluorescent Stains, with Special Reference to Amyloid and Connective Tissues.” *Archives of Pathology* 68 (November):487–98.

Vidarsson, Gestur, Gillian Dekkers, and Theo Rispens. 2014. “IgG Subclasses and Allotypes: From Structure to Effector Functions.” *Frontiers in Immunology* 5 (October). https://doi.org/10.3389/fimmu.2014.00520.

Virchow, Rud. 1854. “Zur Cellulose —Frage.” *Archiv für pathologische Anatomie und Physiologie und für klinische Medicin* 6 (3):416–26. https://doi.org/10.1007/BF02116546.

Westermark, P., K. Sletten, B. Johansson, and G. G. Cornwell. 1990. “Fibril in Senile Systemic Amyloidosis Is Derived from Normal Transthyretin.” *Proceedings of the National Academy of Sciences* 87 (7):2843–45.

Whittaker, E. J. W. 1982. “Crystallography an Introduction for Earth Science (and Other Solid State) Students.” *Crystal Research and Technology* 17 (3):372–372. https://doi.org/10.1002/crat.2170170321.

Williams, David B., and C. Barry Carter. 1996. “The Transmission Electron Microscope.” In *Transmission Electron Microscopy*, 3–17. Springer, Boston, MA. https://link.springer.com/chapter/10.1007/978-1-4757-2519-3_1.

Williams, James C., and Neil Paton. 1978. “CHAPTER 40 - Transmission Electron Microscopy.” In *Systematic Materials Analysis*, edited by J. H. Richardson and R. V. Peterson, 407–75. Academic Press. http://www.sciencedirect.com/science/article/pii/B9780125878043500194.

Wlodawer, Alexander, Wladek Minor, Zbigniew Dauter, and Mariusz Jaskolski. 2008a. “Protein Crystallography for Non-Crystallographers, or How to Get the Best (but Not More) from Published Macromolecular Structures.” *The FEBS Journal* 275 (1):1–21. https://doi.org/10.1111/j.1742-4658.2007.06178.x.

———. 2008b. “Protein Crystallography for Non-Crystallographers, or How to Get the Best (but Not More) from Published Macromolecular Structures.” *FEBS Journal* 275 (1):1–21. https://doi.org/10.1111/j.1742-4658.2007.06178.x.

Wood, E. J. 2004. “Cellular and Molecular Immunology (5th Ed.): Abbas A. K., and Lichtman, A. H.” *Biochemistry and Molecular Biology Education* 32 (1):65–66. https://doi.org/10.1002/bmb.2004.494032019997.

Woody, A-Young M., Richard R. Reisbig, and Robert W. Woody. 1981. “Spectroscopic Studies of Congo Red Binding to RNA Polymerase.” *Biochimica et Biophysica Acta (BBA) - Nucleic Acids and Protein Synthesis* 655 (1):82–88. https://doi.org/10.1016/0005-2787(81)90069-1.

Woolfson, D. N., and D. H. Williams. 1990. “The Influence of Proline Residues on Alpha-Helical Structure.” *FEBS Letters* 277 (1–2):185–88.

Wu, Chun, Zhixiang Wang, Hongxing Lei, Wei Zhang, and Yong Duan. 2007. “Dual Binding Modes of Congo Red to Amyloid Protofibril Surface Observed in Molecular Dynamics Simulations.” *Journal of the American Chemical Society* 129 (5):1225–32. https://doi.org/10.1021/ja0662772.

Yao, Zhi-Xing, and Vassilios Papadopoulos. 2002. “Function of Beta-Amyloid in Cholesterol Transport: A Lead to Neurotoxicity.” *FASEB Journal: Official Publication of the Federation of American Societies for Experimental Biology* 16 (12):1677–79. https://doi.org/10.1096/fj.02-0285fje.

Yifrach, Ofer, and Amnon Horovitz. 1995. “Nested Cooperativity in the ATPase Activity of the Oligomeric Chaperonin GroEL.” *Biochemistry* 34 (16):5303–8. https://doi.org/10.1021/bi00016a001.

Zaros, C., E. Genin, U. Hellman, M. A. Saporta, L. Languille, M. Wadington-Cruz, O. Suhr, M. Misrahi, and V. Planté-Bordeneuve. 2008. “On the Origin of the Transthyretin Val30Met Familial Amyloid Polyneuropathy.” *Annals of Human Genetics* 72 (Pt 4):478–84. https://doi.org/10.1111/j.1469-1809.2008.00439.x.

Zhu, X., X. Zhao, W. F. Burkholder, A. Gragerov, C. M. Ogata, M. E. Gottesman, and W. A. Hendrickson. 1996. “Structural Analysis of Substrate Binding by the Molecular Chaperone DnaK.” *Science (New York, N.Y.)* 272 (5268):1606–14.

Zhu, Ying, Li-Na Zhu, Rui Guo, Heng-Jun Cui, Sheng Ye, and Qun Fang. 2014. “Nanoliter-Scale Protein Crystallization and Screening with a Microfluidic Droplet Robot.” *Scientific Reports* 4 (May). https://doi.org/10.1038/srep05046.

Zou, Kun, Jian-Sheng Gong, Katsuhiko Yanagisawa, and Makoto Michikawa. 2002. “A Novel Function of Monomeric Amyloid Beta-Protein Serving as an Antioxidant Molecule against Metal-Induced Oxidative Damage.” *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 22 (12):4833–41.

Zwanzig, R, A Szabo, and B Bagchi. 1992. “Levinthal’s Paradox.” *Proceedings of the National Academy of Sciences of the United States of America* 89 (1):20–22.